2300

Department of Energy
Ohio Field Office
Fernald Area Office
P. O. Box 538705
Cincinnati, Ohio 45253-8705
(513) 648-3155

JUN 11 1999

Mr. James A. Saric, Remedial Project Manager
U.S. Environmental Protection Agency
Region V-SRF-5J
77 West Jackson Boulevard
Chicago, Illinois 60604-3590

DOE-0729-99

Mr. Tom Schneider, Project Manager
Ohio Environmental Protection Agency
401 East 5th Street
Dayton, Ohio 45402-2911

Dear Mr. Saric and Mr. Schneider:

TRANSMITTAL OF RESPONSES TO THE U.S. ENVIRONMENTAL PROTECTION AGENCY COMMENTS ON THE JANUARY 1999 COMPARABILITY STUDY, CERTIFICATION COMPARABILITY STUDY, REAL TIME RADIATION TRACKING SYSTEM APPLICABILITY STUDY, ADDITIONS TO THE USER GUIDELINES MANUAL, AND DATA VALIDATION CHECKLIST

The purpose of this letter is to transmit responses to the U.S. Environmental Protection Agency (U.S. EPA) and Ohio Environmental Protection Agency (OEPA) comments on the January 1999 Comparability Study; January 1999 Certification Comparability Study; January 1999 proposed additions to the User's Manual; 1999 Data Validation Checklist for ASL D data, and Revision 2.0 of the Real Time Radiation Tracking System (RTRAK) Applicability Study.

At the April 13, 1999 Technical Workgroup meeting, U. S. EPA agreed that Department of Energy, Fernald Environmental Management Project (DOE-FEMP) should respond to comments, since many of the comments did not pertain to the use of *in situ* high-purity germanium (HPGe) detectors for certification.

000001

RESPONSES TO U.S. EPA TECHNICAL REVIEW COMMENTS ON THE "RADIATION TRACKING SYSTEM (RTRAK) APPLICABILITY STUDY" (REVISION 2, JANUARY 1999)

FERNALD ENVIRONMENTAL MANAGEMENT PROJECT

GENERAL COMMENTS

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 1

Comment: The calibration of the RTRAK based on HPGe measurements raises several concerns. First, efficiency calibrations were made by developing calibration algorithms to relate static RTRAK measurements to HPGe measurements. Calibration verification with the RTRAK system was made by taking repeated static measurements in the Uranium in Soils Identification Demonstration (USID) area and Drum Bailing Area (DBA) and comparing those measurements to collocated HPGe measurements. Although the algorithm is reasonable for low activities in the USID area, calibration could not be demonstrated in the DBA for elevated activity. In fact, all of the RTRAK results are biased low compared to the collocated HPGe measurements. Because the bias could approach 20 to 50 percent of the actual HPGe results, areas where measurements are two to three times the FRL warrant concern. The primary problem with the calibration algorithm is that the majority of data points correspond to data points for areas indicative of background levels. These low levels, therefore, govern the calibration curve. Additional calibration data should be collected in areas of higher activity.

Second, calibration verification should be conducted and verification demonstrated for areas that have higher activities. Because the RTRAK system is the first step in determining general areas of elevated contamination, as well as areas that have hot spots (two times the FRL, three times the FRL, and so on), calibration and subsequent verification should focus on these high-activity areas. These issues should be addressed and the text revised accordingly.

Response: Static RTRAK calibration measurements are low relative to the HPGe results at the two locations used for verification in the DBA (Table 4-19). The poorer agreement in the DBA compared to that observed in the USID area may be due to different fields of view of the RTRAK and HPGe detectors (HPGe measurements were made using a 1.0-meter detector height) combined with the presence of heterogeneity in the DBA as discussed in Section 4.5.1 of the report. Note also that the HPGe and RTRAK results for radium-226 agree within the limits of the stated uncertainties for both locations in the DBA and the results of thorium-232 agree within those limits for one of the two locations.

The calibration is not "governed" by data points indicative of background levels. Such data points tend to cluster near the origin and the calibration actually is most influenced by the data points at high concentrations. It is unlikely that needed new locations with high activity for use in calibration or verification can be found on the FEMP site. Not all locations with high concentrations of uranium are suitable candidates, because measurements at such locations tend to be affected by severe interferences from thorium-232 and/or radium-226.

In the future, the FEMP will attemp to improve the RTRAK calibration by using a calibration pad that is being constructed at the site. However, DOE-FEMP believes that the current RTRAK calibration equations are sufficiently accurate for continued use of the instrument given that the range of concentrations for which the RTRAK is calibrated extends to more than three times the FRLs for radium-226 and thorium-232 and approaches that range for total uranium. Additionally, DOE-FEMP believes that the RTRAK calibration is sufficiently accurate for ASL A data.

Action: None.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 2

Comment: Calibration of the radiation scanning system (RSS) is similar to that for RTRAK, where a calibration algorithm was developed based on co-located HPGe and static RSS measurements. However, in this case, calibration verification is demonstrated by comparing dynamic RTRAK and RSS survey results. Similarly, this verification was conducted in the USID (background) area and DBA. Although the USID area showed comparatively similar RTRAK and RSS results, the DBA demonstrated effects resulting from heterogeneitys. However, because calibration of the RTRAK system is questionable, comparison of RTRAK to RSS results is also questionable. In this case, RSS calibration verification should have been conducted using static measurements collocated against the HPGe system. These issues should be addressed and the text revised accordingly.

Response: DOE-FEMP believes that the method of RTRAK and RSS calibration is appropriate for ASL A data. Further, because the RSS detection system, signal processing electronics, and gamma spectrometry software are identical to those on the RTRAK, the comparison of RTRAK and RSS data are valid. Larger discrepancies between the two data sets are permissible than would be the case for ASL D data. However, as observed in Tables C-10 through C-12, average absolute differences are still less than 20 percent in the DBA.

Many of the issues concerning calibration of the RTRAK and RSS should be resolvable after the calibration pad is constructed. It is DOE-FEMP's intention to try and establish calibration based upon a suite of NIST traceable sources of known activity distributed within the field of view. Further, DOE-FEMP intends to use the calibration pad to experimentally determine the field of view of both RTRAK and RSS. In summary, by calibrating HPGe, RTRAK, and RSS on the same calibration pad using the same sources, it will not be necessary to verify calibrations by checking RTRAK or RSS against HPGe. Refer also to the response to Original General Comment No. 4.

Action: Calibrate RTRAK and RSS against known standards using the calibration pad per the discussion above.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 3

Comment: Although theoretical detection levels corresponding to area and radionuclide activities have been developed, it has not been demonstrated that these levels are achievable in the field. Based on the data provided in Table 4-19 for the RTRAK system and Tables C-1 through C-3 for the RSS, it is not clear if these instruments can adequately identify contamination hot spots. This issue should be further addressed and the text revised accordingly.

Response: The RTRAK detection limits were estimated in a manner analogous to laboratory detection limits for radiological techniques, using the approach defined by Currie. By definition, detection limits calculated in this way are "theoretical." The detection is an *a priori* estimate based entirely on the precision of the instrument. It does not incorporate either accuracy or factors unique to the material being measured. The detection capability of an instrument in the field requires knowledge of *a posteriori* factors such as interfering radionuclides, which would not be appropriately incorporated into the detection limit. The capability to detect radionuclides at a given level under location-specific conditions must be evaluated at that location.

The detection limit values quoted in Table 4-14 are comparable to the values presented in Table 4-19 for both HPGe and RTRAK. There are some discrepancies between the HPGe and the RTRAK measurements, but these are not near the stated detection limit. The results in Table 4-19 indicate that the instrument can measure contaminant concentrations at levels below FRLs.

Data in Table 4-19 should not be taken as an indication of the RTRAK ability to identify contamination hot spots. RTRAK measurements are low relative to HPGe measurements for locations in the DBA. This probably results from differences in the field of view between static RTRAK and HPGE. Data in Table 4-21 shows the RTRAK ability to detect total uranium, thorium-232, and radium-226 hot spots. Differences between HPGe and RTRAK data in Table 4-21 are discussed on Pages 4-35 and 4-36.

Action: None.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 4

Comment: It is not clear why a 31-centimeter (cm) detector height was used for the HPGe system for calibration purposes. The text states that the 31-cm height produced the best results for calibration purposes. However, because the HPGe detector field of view is almost four times larger than that of the static RTRAK or RSS field of view, it is not clear how the RTRAK and RSS instruments could be used to demonstrate hot spot detection. The text should provide a more adequate explanation for use of the 31-cm HPGe detector height because the RTRAK and RSS systems may be typically used in heterogeneous areas. Otherwise, it appears that the RTRAK and RSS systems should be calibrated against the 15-cm HPGe detector height. This issue should be further considered and the text revised accordingly.

Response: In areas for which radionuclides are uniformly distributed, matching the fields of view of the HPGe or NaI detectors for calibration is not an issue since results should be the same independent of the field of view that is used. However, if the radionuclides are not uniformly distributed, then three geometric factors may result in inconsistencies between HPGe and RTRAK measurements: (1) any difference in the positions of the HPGe and NaI detectors when measurements are made, (2) differences in the sizes of the fields of view of the detectors, and (3) differences in the shapes of the fields of view of the detectors. The NaI detector does not have a circularly shaped field of view. Additionally, the orientation of the NaI detector (which is determined by the orientation of the NaI detector on the RTRAK) relative to any pattern in the distribution of the radionuclides may have some effect on RTRAK results. Areas with a significant amount of heterogeneity in the distribution of radionuclides were not used for calibration of the RTRAK. However, some small amount of heterogeneity will always be present that will affect the agreement between HPGe and RTRAK results due to the three factors listed above. Therefore, achieving the best fit between RTRAK and HPGe results was used as the basis for establishing the HPGe detector height to be used when calibrating the RTRAK. Regressions were carried out using data obtained for HPGe detector heights of 15 cm, 31 cm, and 1 m. The best results were obtained using 31-cm data. Statistically, 31-cm HPGe data provide the best match with the RTRAK results for the data set used for calibration. Therefore, calibration using a 31-cm HPGe detector height is preferred. Uncertainties associated with any less than perfect agreement between HPGe and RTRAK results, including any differences related to fields of view, are reflected in the uncertainties in the parameters in the calibration equations.

With regard to hot spot detection, the identification of hot spots by RTRAK is not a matter of RTRAK calibration, but rather depends upon the activity of the hot spot and its size relative to the size of the RTRAK moving field of view. These issues are discussed in Section 3.3 (Hot Spot Detection) of the User's Manual. Figure 3.3-2 in that section shows the size versus concentration relationship of hot spots for RTRAK detection.

Action: As stated in the response to Original General Comment No. 1, the RTRAK/RSS calibration will be re-evaluated after construction of the calibration pad is complete. This will include assessing the appropriate HPGe height for calibration measurements. However, it is possible that use of a calibration pad to calibrate RTRAK and RSS may eliminate the need to calibrate RTRAK/RSS by comparison the HPGe measurements.

SPECIFIC COMMENTS

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 2.1 Page #: 2-1 Line #: 7
Original Specific Comment #: 1

Comment: The text states that Figure 2-1 shows the location of the USID and South Field study areas. In addition, line 19 on page 2-4 states that Figure 2-1 shows the USID study area. However, Figure 2-1 shows the "Southfield Area" but not the USID area. The figure should be corrected to show the USID area.

Response: Agree with comment.

Action: The next revision of the RTRAK report will include a revised Figure 2-1 figure or an additional figure to show the location of the USID area.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 4.1.1 Page #: 4-17 Lines #: 1 through 4
Original Specific Comment #: 2

Comment: These lines present equations 6 through 8. Each of these equations includes the term "t^2" in the denominator of the right-hand side. The text should be revised to define this term.

Response: Agree with comment; the "t^2" term is not defined. Additionally, the equations are in error in that the t^2 factor should be in the denominator inside the square root.

Action: The equations will be revised for the next revision of the RTRAK report, and the "t^2" term will be defined.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 4.5 Page #: 4-31 Line #: 19
Original Specific Comment #: 3

Comment: This section compares RTRAK measurements to HPGe measurements. RTRAK results should also be compared directly to the corresponding laboratory analytical results. This direct comparison would minimize the propagation of error in the measurements and calculations. This same comment applies to Section C.3, which describes the RSS.

Response: While it may be desirable to make comparisons directly with laboratory data, there is a shortage of data to allow this comparison to be made. Because discrete samples are so small in comparison to the field of view of the RTRAK, comparisons with individual samples would be highly influenced by heterogeneity. An effective comparison would require collecting samples in a weighted pattern similar to that used for the HPGe comparability study. The pattern for the HPGe was based on models used to define the HPGe behavior, and similar models have not been developed for the geometry of the RTRAK detector. In this regard, note that the HPGe comparability study demonstrates good comparability of HPGe measurements with laboratory analytical measurements. Thus, for RTRAK, direct calibration or verification of calibration with laboratory data is not necessary. What is necessary is an effective calibration against HPGe that produces data of a requisite quality for its intended use. The current RTRAK calibration meets that requirement.

Action: None.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 4.5.2 Page #: 4-36 Line #: 17
Original Specific Comment #: 4

Comment: The text states that Table 4-22 compares certain HPGe results with corresponding RTRAK results. However, Table 4-22 is omitted from both this section and the table of contents. Either the text should be revised to include Table 4-22, or the citation in the text should be corrected.

Response: Agree with comment. Table 4-22 was inadvertently omitted from the report.

Action: Table 4-22 will be included in the next revision of the report.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 5.2.1 Page #: 5-4 Line #: 27
Original Specific Comment #: 5
Comment: The text states that some results are summarized in Table 5-3. This citation should be corrected to read "Table 5-2."

Response: Agree with comment. The citation should read Table 5-2.

Action: The citation will be corrected in the next revision of the RTRAK report.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 6.0 Page #: 6-4 Lines #: 4 and 5
Original Specific Comment #: 6
Comment: The bold-faced text of this conclusion is incomplete; therefore, the text is unclear. The text should be revised to include the missing words for clarification.

Response: The bold-faced statement should read: "The current calibration equations provide good agreement with HPGe except at low uranium concentrations."

Action: The text will be corrected in the next revision of the RTRAK report.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: A.3.1 Page #: A-3 Line #: 25
Original Specific Comment #: 7
Comment: The text states that pulse height is proportional to gamma ray energy. Actually, pulse height in a multichannel analyzer is proportional to gamma ray flux (number of gamma rays), but the channel on which the pulse appears is proportional to the individual gamma ray energy (that is, its frequency or inverse wavelength). The text should be revised accordingly.

Response: The text is correct as currently written. The signal (i.e., pulse height) from a gamma detector is proportional to the energy of the gamma ray. The commentor is confusing **pulse height** with the concept of spectral **peak height**, which is proportional to the gamma-ray flux.

Action: None.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: A.4.1 Page #: A-8 Line #: 21
Original Specific Comment #: 8
Comment: The text describes intermittent limitations on the quality of global positioning system (GPS) data. It would be useful to note that quality is generally a function of the number of satellites with an unobstructed line of sight to the receiving antenna. This section should be revised to note that the number of satellites from which data are actually received, as well as other parameters such as discontinuities in apparent position, are used for QC purposes in assessing real-time data.

Response: The purpose of Appendix A was to provide an overview of the RTRAK/RSS and ancillary systems. Details such as those that the reviewer recommends including in Appendix A are included in operating procedures and Quality Control documents. They were not put into the RTRAK Applicability Study because they were not a key element in the RTRAK/RSS method validation study.

Action: Appendix A will be revised to include the details requested.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: C.1 Page #: C-1 Line #: 21
Original Specific Comment #: 9
Comment: The text states that the RSS is essentially identical in operation to the RTRAK, with a gain of 5.84 kiloelectronvolts (keV) per channel. However, line 17 on Page A-4 states that the RTRAK has a gain of 5.85 keV per channel. Because input data (channel number and peak energy) are the same for the RSS and the RTRAK, the same calculation result, including rounding off, should be presented for both systems. The text should be revised accordingly.

Response: Agree with comment. There is a discrepancy in the text. The keV/channel factor should be 5.84 in both locations.

Action: The text will be corrected in the next revision of the RTRAK report.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: C.4 Page #: C-13 Line #: 19
Original Specific Comment #: 10
Comment: The text concludes that strong statistical evidence indicates that the RTRAK and RSS measurements are equivalent. Because of the irregularities of the results from the USID area, which are discussed in Section C.4, it is more accurate to state that the two systems produce equivalent results at levels that reasonably exceed the minimum detection concentration. The text should be revised accordingly.

Response: The authors of the RTRAK Applicability Study believe that the statistical evidence that the two systems produce equivalent results is strong. Sixty separate t-tests were performed, resulting in forty four decisions that the RSS and RTRAK segment means were equivalent at the 95 percent confidence level. However, the reviewer's comment is correct. It would be more accurate to state that the two systems produce equivalent results at radionuclide concentrations that reasonably exceed the minimum detectable concentration.

Action: The text will be revised to include this statement in the next revision of the RTRAK Report.

RESPONSES TO U.S. EPA TECHNICAL REVIEW COMMENTS ON THE "COMPARABILITY OF IN SITU GAMMA SPECTROMETRY AND LABORATORY DATA" (JANUARY 1999, REVISION 1)

FERNALD ENVIRONMENTAL MANAGEMENT PROJECT

GENERAL COMMENTS

Commenting Organization: U.S. EPA — Commentor: Saric
Section #: Not applicable (NA) — Page #: NA — Line #: NA
Original General Comment #: 1

Comment: The comparability study involves a number of comparisons of laboratory methods, especially for analysis of total uranium and radium 226. Work has been started on a "Multi-Agency Radiation Laboratory Protocols (MARLAP) Manual" involving the Department of Energy (DOE), U.S. Environmental Protection Agency (U.S. EPA), and five other federal agencies. The method comparison data and other information generated while resolving data discrepancies should be provided to the MARLAP group to support development of the MARLAP Manual.

Response: Agree with comment.

Action: This is a good suggestion. If MARLAP is interested, we will send relevant information to MARLAP via DOE's Environmental Measurements Laboratory.

Commenting Organization: U.S. EPA — Commentor: Saric
Section #: NA — Page #: NA — Line #: NA
Original General Comment #: 2

Comment: In situ gamma spectrometry using a high-purity germanium (HPGe) detector has not yet been accepted as a source of definitive data, especially for certifying that soil in an area meets final remediation levels (FRL) for primary radionuclides. As discussed in original specific comment 4 below, the radium 226 HPGe results appear to have a definite low bias, which is unacceptable for making final certification decisions. In addition, as discussed in comments on the draft data validation checklist for validating HPGe detector measurements to analytical support level (ASL) D, the current calibration method that uses multiple readings of a point source that are mathematically manipulated is quite different from the analytical method that measures an overall, relatively uniform source. Therefore, the use of data from the HPGe detector for final certification purposes is premature.

Response: DOE-FEMP agrees that radium-226 has a definite low bias relative to laboratory data. However, DOE-FEMP believes that sufficient data have been accumulated to demonstrate that this low bias can be adequately compensated for. In addition, administrative controls such as those described in proposed Section 3.7 to the User's Manual can be implemented that default to physical samples when radium-226 data does not meet certain pre-specified criteria.

The method of calibration of in-situ gamma spectrometry involving a mathematical model combined with measurements of certified standards has been consistently demonstrated to be valid over the past 27 years since the approach was first published by Beck (1972).

This method of calibration is routinely used by all in-situ gamma spectrometry groups, including the U.S. EPA enforcement gamma spectrometry group. In addition, it follows international guidance (ICRU 53) and is consistent with a draft ANSI standard being prepared. Just because the calibration method is different from that used for laboratory gamma spectrometry is not grounds for rejection of the method. Accordingly, DOE-FEMP disagrees with this portion of the comment.

Action: None.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 3

Comment: The comparability study demonstrates reasonable agreement between HPGe detector measurements and laboratory results for uranium and thorium 232. However, some concerns still exist regarding HPGe detector measurements of radium 226. In this comparability study, a new correction algorithm is provided that differs from the one proposed in the user's manual for the real-time instruments. In fact, the new algorithm results in corrected radium 226 concentrations that are markedly lower than those using the user's manual method. In addition, little information exists to justify the use of the revised correction algorithm. The text states that Section 8.0 of the comparability study presents a validation of the correction algorithm. However, Section 8.0 does not provide adequate information to validate the correction algorithm. To begin with, none of the validation results apply to radium 226 concentrations near the FRL of 1.7 picocuries per gram (pCi/g). Table 4 in Appendix J shows area-average laboratory analysis results from 0.70 to 0.99 pCi/g and in situ HPGe results from 0.58 to 0.87 pCi/g. Furthermore, for the majority, if not all, of the data in the validation study, comparability between data points can be demonstrated without the use of a correction algorithm. Therefore, the new correction algorithm is not appropriately tested and its use is not validated. The text should be revised to further address these issues.

Response: The correction algorithm developed in Section 7.6.2 is different than the one in the User's Manual. The one in the User's Manual is based only upon laboratory alpha spectrometry data, while the one in Section 7.6.2 is based upon the average of laboratory alpha and gamma spectrometry data. Because laboratory gamma spectrometry yielded concentrations of radium lower than those from alpha spectrometry, the new correction factor results in lower corrected radium-226 concentrations than those calculated by the algorithm in the User's Manual. DOE-FEMP believes the new algorithm is superior to the old, because it is based upon the average of multiple laboratory analytical methods. Moreover, use of a correction factor based upon multiple laboratory analytical methods is consistent with the rest of the Comparability Study in that comparability is also based upon multiple analytical methods.

Clearly, as shown in Figure 7-2, at low concentrations (less than 1.0 pCi/g) there is not much difference between HPGe and laboratory data. Although HPGe radium-226 is biased low, correction is not necessary to meet comparability guidelines proposed in this document. The situation is markedly different for radium-226 concentrations increasingly larger than 1.0 pCi/g. The difference between HPGe and laboratory concentrations progressively increases as radium-226 progressively exceeds 1.0 pCi/g. Correction factors are necessary for such data.

Finally, DOE-FEMP disagrees with the a portion of the comment concerning Section 8.0. The closeness of the corrected dry weight radium-226 concentration (1.42 pCi/g in Table 8-5B is virtually identical to the average of laboratory alpha and gamma spectrometry data (1.40 pCi/g). The average of the laboratory data (1.40pCi/g) is sufficiently close to the FRL for radium-226 (1.70 pCi/g) to establish the validity of the correction algorithm at the FRL.

Action: None.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 4

Comment: The comparability study raises a general concern regarding HPGe measurements and data quality. During physical sampling activities, strict laboratory quality control (QC) must be adhered to. One element associated with laboratory QC involves measurement of spiked samples and their duplicates. In this case, the measured value can be directly compared to a known value. However, this is not the case with HPGe measurements. In situ gamma spectrometry can only be used to approximate the true value. Therefore, a flaw exists with regard to HPGe QC criteria. According to the text in Section 9.0, ASL D requires HPGe values to be within 20 percent of the laboratory values. However, because the laboratory values are allowed to be within 20 percent of the true (known) values, HPGe results are allowed to potentially be within 40 percent of the known values. Because a 40 percent allowance is unacceptable, one of the following two solutions should be considered: (1) lowering the QC acceptance criterion for laboratory samples and HPGe measurements (combined) so that the total propagated difference is less than 20 percent or (2) ensure that the sum of the average relative percent differences (RPD) for physical samples and matrix spikes, and for HPGe measurements and physical samples, (the total RPD) is less than 20 percent.

Response: DOE-FEMP recognizes that the lack of certain QC elements such as laboratory control standards (LCS) represents a difference inherent in in-situ gamma spectrometry relative to laboratory analyses. This does not mean, however, that in-situ gamma spectrometry only approximates the true value. The HPGe system depends upon NIST traceable sources of known activity in the calibration process to yield a true value. In this regard, it must be recognized that the detector calibration is based on fluence rate which is itself traceable to those NIST certified standards. To demonstrate comparability HPGe measurements must be compared against laboratory data for real samples collected from within the field of view of the HPGe detector at a given location. If there are bias errors in the laboratory data, the HPGe data could be closer to or further from true values than is believed.

These limitations are thoroughly discussed in Section 4.2 and are the reason why multiple measurements involving different analytical techniques were utilized. DOE-FEMP believes that the average of multiple analytical techniques represents the best approach to obtaining "true" laboratory data and is to be preferred over error propagated methods (such as approach 1 in the comment above) in which key uncertainties may have to be grossly estimated.

Comparability is simply a check on the results of two independent measurements, typically for the purposes of detecting any bias in a new method and to check the reasonableness of the respective uncertainty estimates. The presence of bias in HPGe data would be indicated if the 20 percent criterion were exceeded. As discussed in Section 4.2, the reason multiple measurements involving different analytical techniques were utilized was to minimize any possible bias in the laboratory data. Regarding proposed solution 1, above, it is not appropriate to propagate error across independent analysis methods. Rather, errors should only be propagated within a method along the fundamental steps that make it up. The 20 percent RPD criterion does not imply that new (HPGe) method is potentially 40 percent off the true value simply because laboratory methods also apply a 20 percent comparability criterion. In these comparisons, the established method (in this case, the laboratory method) must be assumed to be unbiased, otherwise it would not be suitable for assessing bias in the new method. Having made a reasonable effort to root out bias in the new method by comparing it to a method established to be unbiased, any estimate of the possible variance of a measurement from the true value can only be made on the basis of a propagation of error within the new method.

Approach 2 is not a valid approach because matrix spikes are not used in laboratory gamma spectrometry. Moreover, as in approach 1, it suggests inappropriately that RPDs be used in an additive fashion rather than as a threshold for detecting bias. In short, if a laboratory method is within the 20 percent RPD criterion of a known sample (e.g., matrix spike), it can be assumed to be unbiased. The same 20 percent RPD criterion can then be used to compare a new method to the laboratory method to detect bias. In general, matrix spikes are unreliable QC parameters to be used in any type of error propagation or error summation. The commentor is referred to a discussion of problems associated with the use of matrix spikes in the book "Environmental Laboratory Data Evaluation," Genium Publishing Corporation, pp 218-220, 1996. For many of the reasons discussed the this book, the FEMP decided to use the average of multiple analytical methods as the best approach to obtaining the most accurate laboratory data.

Action: None.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 5

Comment: In various places, the text states that the total propagated uncertainty (TPU) of laboratory data is higher than the TPU for HPGe measurements. The text states that this results from the fact that available HPGe data only account for counting errors and that the HPGe system will register many more counts. However, because the HPGe system relies on comparison to laboratory data to approximate the "true" value, errors associated with laboratory measurement should be included when calculating HPGe measurement errors. Unless complete TPUs for the HPGe system are provided, discussion on measurement error between the two methods is useless. If complete TPUs for the HPGe system are not provided, only the counting error associated with laboratory measurements should be included.

Response: The HPGe system does not depend upon comparison with laboratory data to approximate a true value. The HPGe system depends upon the calibration method to yield an accurate value. There is no need to incorporate laboratory errors into the TPU from HPGe. As discussed in Sections 2.6 and 2.7, standard deviations of 250 measurements at the FCS over a years time can be interpreted to represent the long-term system total uncertainty.

DOE-FEMP agrees that a comparison between laboratory and HPGe errors should compare counting errors to counting errors. The convention at the FEMP is for all laboratory uncertainties to be reported as TPUs. Generally about 95 percent or more of the reported laboratory TPUs are due to counting error and less than 5 percent of the reported TPUs are due to other contributions. Thus, for all practical purposes the reported laboratory TPUs are approximately equal to counting errors.

Action: None.

SPECIFIC COMMENTS

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 2.8 Page #: 2-11 Line #: 4
Original Specific Comment #: 1
Comment: The text states that some "statistically significant differences" exist between concentrations measured by different detectors. Table 2-5 shows some pairs of data sets as differing at the 95 percent confidence level. The text should be revised to identify the statistical test used to support the conclusion that some pairs of data differ significantly.

Response: Agree with comment.

Action: The text on Page 2-11 will be revised to indicate that two sample "F" tests for equality of variances and two sample "t" tests for equality of means were utilized.

Commenting Organization: U.S. EPA Commentor: Saric
Figures #: 3-2 and 3-3 Page #: NA Line #: NA
Original Specific Comment #: 2
Comment: Figures 3-2 and 3-3 show additional study locations. To increase the useability of these figures, Figures 3-2 and 3-3 should be revised to include the symbols for low, medium, and high concentration areas that are used in Figure 3-1. Also, Figure 3-1, which shows the entire facility, should be revised to indicate the general locations of the detailed maps in Figures 3-2 and 3-3. These modifications would help the reader relate the study areas to the overall pattern of contamination within the facility.

Response: Agree with comment.

Action: The figures will be revised in accordance with the comment.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 5.3 Page #: 5-2 Line #: 10
Original Specific Comment #: 3

Comment: Section 5.3 discusses total uranium analytical results, and the accompanying tables and appendixes provide data for areas 0 through 10. However, the text of Section 3.2.1 discusses collection of data from areas 12 through 19 to supplement data already available from areas 1 through 10. Either the tables in Section 5 and associated appendixes should be revised to include data for areas 12 through 19, or the omission of these data should be justified in the text.

Response: Clarifying text will be added in Section 5.1 indicating that total uranium analyses were only performed for Areas PBC-01 through PPB-10 and PBC-18 and 19. PBC-18 and 19 were analyzed subsequent to the July 1997 Comparability Study to extend HPGe comparability for uranium to WAC concentrations. These data were reported in the September 1997 addendum to the July 1997 Comparability Study. PBC-12 through 17 were analyzed subsequent to the July 1997 Comparability Study to infill radium-226 data gaps. These results were reported in the September 1997 addendum to the Comparability Study.

Action: Clarifying text will be added in Section 5.1.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 7.4 Page #: 7-5 Line #: 5
Original Specific Comment #: 4

Comment: This section discusses the data in Tables 7-3 and 7-4 and compares results from laboratory analyses and HPGe detector measurements for radium 226. However, the text does not note that the HPGe detector results are lower than the results from both laboratory methods for 6 of the 10 areas in Table 7-3 and 12 of the 16 areas in Table 7-4. If all three methods give equal results, then the HPGe detector should yield the lowest results one-third of the time. Using the binomial distribution, the probability of the actual results occurring by chance alone is 0.0197 for Table 7-3 data and 0.0001 for Table 7-4 data. These probabilities indicate a systematic difference between the HPGe and laboratory results for radium 226, especially at the 1-meter height (Table 7-4). The high coefficient of variation for the radium 226 results, which results from many of the data points being near the minimum detectable activity, indicates that parametric tests for bias among the analytical methods have very low statistical power. Therefore, nonparametric tests, such as the binomial test discussed above, are more appropriate for comparisons of laboratory results and HPGe detector measurements. The text should be revised to discuss the observed low bias in the HPGe detector results for radium 226, including how the low bias is related to the correction algorithm.

Response: The text on Page 7-6, Lines 13-17, indicates that HPGe radium-226 results are biased low to laboratory data and why they are biased low. The text on Page 7-7, Lines 9-12, reiterates the point.

DOE-FEMP believes that Section 7.6 adequately relates the low bias in the HPGe data to the correction algorithm.

Action: None.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 7.4.2 Page #: 7-7 Line #: 3
Original Specific Comment #: 5
Comment: Based on the data provided in Tables 7-5 and 7-6, it appears that HPGe and laboratory data for radium 226 are less comparable for HPGe measurements collected at the 1-meter detector height. Therefore, it is not clear why the 1-meter HPGe detector height was selected for acquiring additional data points. The text should be revised to clarify this issue.

Response: The reason for only acquiring 1 meter data has historical roots. A conclusion of the July 1997 Comparability Study was that differences between 31 cm and 100 cm detector height data were minimal. Further, at that time the assumption was that the HPGe would be used only at a one meter detector height. When additional data were collected for radium, then, only a 1.0 meter detector height was used, and time and money were spent in collecting a.m. and p.m. measurements as indicated in Section 7.3, Lines 12-23.

Action: Clarifying text will be added to Section 7.3.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 7.6.2 Page #: 7-10 Line #: 9
Original Specific Comment #: 6
Comment: The correction algorithm discussed in Section 7.6.2 is based on all available radium data. The correction algorithm should be validated by collecting new HPGe data, calculating HPGe results using this algorithm, and comparing the results to analytical results from physical samples.

Response: The correction algorithm is not based upon data from the field quality control station. Independent verification of the algorithm is demonstrated in Section 8.4.1, Page 8-4, Lines 19-21. Additionally, independent verification of the algorithm is provided by comparing HPGe measurements with laboratory measurements of physical samples collected in CU's from four areas and reported in the document entitled "Comparability of In-Situ Gamma Spectrometry and Laboratory Data and Decisions for Certification Units."

Action: None.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 7.6.2 Page #: 7-10 Line #: 14
Original Specific Comment #: 7
Comment: The text states that the two-method average of laboratory gamma and alpha analytical results was used because these averages most likely represent the true radium concentrations. However, it is not clear what this statement is based on. In addition, the laboratory data should be supported by QC measurements such as matrix spike and matrix spike duplicate results. The method of measurement that yields results closest to the spiked sample concentrations is usually more appropriate. If multiple spikes are used that yield mixed results between the two methods, then an average of the results from the two methods may be more appropriately used. However, such QC information is not provided in the text. The text should be revised to include additional justification for selection of the two-method average.

Response: Disagree with comment. DOE-FEMP believes that the discussions in Section 4.3.1 and in Table 4-1 adequately justify the use of the average of multiple methods as being more appropriate than a given single method. Also, note that matrix spikes and matrix spike duplicates are usually not performed in laboratory gamma spectrometry measurements. See also the response to Original General Comment No. 4 concerning matrix spikes.

Action: None.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 7.6.2 Page #: 7-11 Line #: 8
Original Specific Comment #: 8
Comment: The text states that "3.35 - 2.0 = 1.28." The text should be revised to correct this arithmetic or transcription error. The correction factors should then be checked to verify that no such errors exist in the data sets used for calculations and in the calculations themselves.

Response: All spreadsheet calculations and data indicate that "3.35" and "2.0" are correct and that the "1.28" is either a typo or a transcription error not caught in editing.

Action: "1.28" will be changed to "1.35" in Line 8 on Page 7-11.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 9.4 Page #: 9-4 Line #: 18
Original Specific Comment #: 9
Comment: The text presents an equation for calculating the minimum detectable concentration of radionuclides. The text should be revised to cite the source for this equation and to justify its applicability to HPGe measurements.

Response: MARSSIM, Page 6-34, Equation 6-7 gives the MDC as $MDC=C(3.0+4.65\sqrt{B})$. This is equivalent to the MDC equation on Line 18. However, the factor "2.71" is generally used by most radiochemists instead of "3.0," and is the factor used in commercial software such as GammaVision.

Action: The MARSSIM document will be referenced for the MDC equation.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 9.6 Page #: 9-6 Line #: 22
Original Specific Comment #: 10
Comment: This section concludes that in situ HPGe measurements produce definitive (ASL D) data. However, although the calibration technique produces apparently acceptable results in the studies detailed here, it is quite different from the measurement technique. Therefore, the calibration technique may not be sufficiently robust to produce acceptable results under other circumstances such as heterogeneous distribution of radionuclides. In addition, radium 226 HPGe results are biased low, which is not acceptable for certification purposes. The text should be revised to address these issues.

Response: DOE-FEMP believes that the low radium-226 bias is adequately compensated for by using the correction algorithm. Additionally, the strategy for HPGe certification

measurements proposed in Section 3.7 of the User's Manual has a built in "conservatism" factor that defaults to physical samples when radon correction factors are suspect.

DOE-FEMP disagrees with the comment regarding the robustness of calibration. Circumstances such as heterogeneity are not a factor in calibration but of data interpretation. This is discussed in the User's Manual in Section 5.5 entitled "heterogeneity." See also the response to Original General Comment No. 2 regarding HPGe calibration.

Action: None.

Commenting Organization: U.S. EPA Commentor: Saric
Appendixes #: A, B, and C Page #: NA Line #: NA
Original Specific Comment #: 11
Comment: These three appendixes contain many pages of data. However, the pages of these appendixes are not numbered. The appendixes should be revised to include page numbers as in Appendix D.

Response: Agree with comment.

Action: Pages in Appendices A, B, and C will be numbered.

Commenting Organization: U.S. EPA Commentor: Saric
Appendix #: G Page #: G-3 Line #: 19
Original Specific Comment #: 12
Comment: The text states that Attachment 3 includes an example report generated by the GammaVision software. However, the appendix contains only two attachments, neither of which is a report generated by the software. Either the GammaVision report should be added or the text should be modified.

Response: Attachment 3 (not included) and Table F-2 in Appendix F were redundant. Attachment 3 was subsequently eliminated in the report, but the text was not changed.

Action: The reference to Attachment 3 on Line 19 of Page G-3 will be dropped.

Commenting Organization: U.S. EPA Commentor: Saric
Appendix #: G Page #: G-4 Line #: 19
Original Specific Comment #: 13
Comment: The equation in Line 19 defines the five-point weighted average used to smooth the data. However, the equation does not include a term for the original (central) data point. The equation should be revised to include a term for the original data point. Also, the coefficient in the denominator is 16, but the sum of the coefficients in the numerator is only 12. The equation should be corrected so that the sums of the coefficients are equal in the numerator and denominator.

Response: Agree with comment. The correct form of the equation is:

$$S_i=((O_{i+2})+(4O_{i+1})+(6O_i)+(4O_{i-1})+(O_{i-2}))/16$$

where: S_i = smoothed data in channel i
O_i = original data in channel i

Action: The equation on Line 19 on Page G-4 will be corrected as shown above.

Commenting Organization: U.S. EPA Commentor: Saric
Appendix #: J Page #: Table 3 Line #: NA
Original Specific Comment #: 14
Comment: The text states that 1.0-foot and 1.0-meter height HPGe measurements for radium 226 demonstrate a significant difference at location 10. However, Table 3 indicates that no such difference apparently exists. Either the text or the table should be revised to resolve this discrepancy.

Response: The text is in error.

Action: The text will be revised to resolve the discrepancy noted in the comment.

RESPONSES TO U.S. EPA TECHNICAL REVIEW COMMENTS ON THE "COMPARABILITY OF IN-SITU GAMMA SPECTROMETRY AND LABORATORY DATA AND DECISIONS FOR CERTIFICATION UNITS" (JANUARY 1999, DRAFT REVISION 0)

FERNALD ENVIRONMENTAL MANAGEMENT PROJECT

GENERAL COMMENTS

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 1

Comment: As detailed in comments below and in comments on the "Comparability of In Situ Gamma Spectrometry and Laboratory Data, Revision 1," sufficient evidence exists to indicate that the HPGe system is unacceptable for certification of radium 226 measurements because radium 226 measurements taken using the HPGe detector are biased low. In addition, insufficient evidence exists to fully support the use of the HPGe system for certification of the uranium 238 (total uranium) and thorium 232 measurements.

Response: DOE-FEMP disagrees with the comment. The closeness of total uranium concentrations as measured by both laboratory and HPGe methods in four certification areas indicate that either method could be employed without loss of confidence in the data quality. The same statement holds for thorium-232.

DOE-FEMP agrees that radium-226 HPGe measurements are biased low relative to laboratory data because of radon-222 disequilibrium. However, empirical correction factors adequately adjust for such disequilibrium a majority of the time. Further, in the proposed strategy presented in Section 3.7 of the User's Manual for the usage of HPGe in certification, a default to physical sampling is mandated when certain radium-226 correction criteria are not met. This builds a strong degree of conservatism into the certification strategy.

Action: None.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 2

Comment: It is not clear if unweighted laboratory data were used in this comparability study for certification units. Previous comparability studies demonstrate to a reasonable degree that laboratory data and HPGe measurements are in general agreement for areas that demonstrate homogeneous contamination. Therefore, it should not matter if the data are weighted or unweighted. However, only unweighted laboratory results should be used to compare certification decisions based on HPGe measurements to laboratory data. The fact that HPGe measurements are intended to be taken in areas of homogeneous contamination further supports the use of unweighted laboratory results for comparison purposes. Either the report should be revised to specifically state that unweighted laboratory results were used, or the study should show comparability between laboratory data and HPGe measurements using unweighted results.

Response: No weighting factors were used. With the exception of A1PI all physical samples and HPGe measurements were co-located. The same statistical calculations to determine whether or not a CU passed or failed certification were used for both physical samples and HPGe measurements. These are presented briefly in Section 2.2.1 and in detail in the Sitewide Excavation Plan (SEP).

Action: Section 2.2.1 will be clarified to indicate that weighting factors were not used in the statistical evaluation of laboratory data.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 3

Comment: The association between validated results and certification decisions is not fully clear with regard to duplicate measurements. For example, HPGe results for Area 1, Phase II (A1PII), indicated that two of the six certification units would have estimated (J) data for radium 226 because the RPD for duplicate measurements exceed 35 percent. For A1PII, the results for these two data points indicate that radium 226 concentrations are below the FRL of 1.7 pCi/g, and have a 1.06 pCi/g 95 percent upper confidence limit (UCL) for S1-08 and a 1.03 pCi/g 95 percent UCL for S1-10. However, it is not clear what impact these estimated results would have on decision-making if they are close to the associated FRL. Further, because the intent of duplicate measurements is to demonstrate instrument precision, failure to meet duplicate measurement criteria results in questionable data. Finally, if consecutive duplicate measurement criteria (that is a 20 percent RPD) were instituted, four of the six certification units would have estimated data for radium 226. This issue should be discussed further in Real-Time Workgroup meetings and the text revised accordingly.

Response: DOE-FEMP agrees that the issue regarding duplicate measurements could to be discussed at a future Technical Workgroup meeting. Basically, there are two options for duplicate measurements: back-to-back and non-consecutive. Back-to-back measurements typically provide information on detector drift and have a 20 percent RPD as noted in the above comment. Nonconsecutive measurements, particularly those taken on different days, may have a major environmental component; especially for radium-226. This component can provide considerable information relative to the usability of HPGe data from an environmental effects perspective that is not obtainable from back-to-back duplicates. This has been factored into the certification strategy for HPGe measurements proposed in Section 3.7 of the User's Manual.

Also, acceptance criteria have been defined for usability of non-consecutive duplicates (per the proposed ASL D Data Validation Checklist) to account for the additional environmental component that is not present in back-to-back duplicates.

Action: Discuss issue of measurement duplicates at a future Technology Workgroup meeting and revise User's Manual and Data Validation Checklist in appropriate sections as applicable.

SPECIFIC COMMENTS

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 2.1 Page #: 2-1 Line #: 27
Original Specific Comment #: 1

Comment: The correction factor for radium 226 discussed in Section 2.1 is different from the correction factor provided in the user's manual for the real-time instruments. According to the creators of the original correction factor, the algorithm is appropriate for correcting HPGe results and additional verification should be performed. However, a new correction factor has now been developed. Past data that used or was used to develop the original correction factor should be assessed to ascertain if comparability is preserved, and independent testing of this new correction factor should be conducted for verification purposes. The text should be revised accordingly.

Response: A1PI certification comparability data reported to the USEPA in the July 1997 Comparability Study were not radon disequilibrium corrected. The original radium correction factors were developed subsequent to the submission of that report. All of the radium-226 data in this study have been reported for the first time on a radium corrected basis (using the new correction algorithm), and the comparison of corrected HPGe data with laboratory data serves as independent verification. Consequently, there is no need to test past and present comparability.

Action: None.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 2.2.1 Page #: 2-3 Line #: 10 through 18
Original Specific Comment #: 2

Comment: The text presents three cases for passing or failing certification. However, the three cases are not complete (do not include all possibilities), and some cases are logically implied but impossible, such as the average, but not the UCL on the average, exceeding the FRL. These cases should be revised to reflect the following: (1) the UCL is less than the FRL, and the maximum result is less than the hot-spot criterion; (2) the UCL is greater than the FRL, and maximum results are irrelevant; and (3) the maximum result is greater than the hot-spot criterion, and UCL results are irrelevant.

Response: The intent of Lines 10 through 18, Pages 2-3 was not to present a comprehensive discussion of all possible cases for passing or failing certification; simply the ones most relevant to this report. That is why the reader is referred to the SEP for more information.

Action: One additional case will be added to Page 2-3, and that is the case where the maximum concentration of any COC is greater than the hot spot criterion.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 3.1 Page #: 3-1 Line #: 16
Original Specific Comment #: 3

Comment: The text states that nine certification unit failures of FRL for thorium 232 result from artifacts of the laboratory analytical procedure. However, the text does not discuss how many other thorium analytical results are unreliable and should also be disregarded. In the absence of more specific data, it must be assumed that all associated thorium 232 results collected in Area 1, Phase I (A1PI), are doubtful. Therefore, very few reliable data sets remain for that radionuclide. The text should be revised to address this issue.

Response: The reliability of thorium-232 certification data acquired by gamma spectrometry from off-site laboratories was a topic of numerous discussions and memoranda between DOE-FEMP, OEPA, and the U.S. EPA. The Th-232 data in A1PI were believed by FEMP personnel to be biased high due to the method of using a single energy gamma peak as the basis for calculation as opposed to a weighted average of multiple gamma energy peaks. Nonetheless, this was difficult to prove conclusively. Therefore, the A1PI data were included in this document. If they are biased high, they represent a worse case scenario relative to comparison with HPGe data.

Action: The text on Page 3-1 will be clarified to reflect the above issues.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 3.2 Page #: 3-3 Line #: 15
Original Specific Comment #: 4

Comment: The text states that none of the certification units in A1PI failed for thorium 232 based on alpha spectrometry data. However, these data do not appear in the report. These data should be included for independent calculation purposes.

Response: The decision was made not to include alpha spectrometry data, because subsequent to A1PI DOE-FEMP committed to the U.S. EPA and OEPA to only use gamma spectrometry data for certification measurements. Thus, no other alpha spectrometry data have been acquired for certification area samples after A1PI. The reader can review the A1PI Certification Report for all details relevant to alpha spectrometry data used for certification in A1PI.

Action: Text will be added to Page 3-3 directing the interested reader to the A1PI Certification Report for all details pertaining to alpha spectrometry data.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 3.4 Page #: 3-5 Line #: 3
Original Specific Comment #: 5

Comment: The conclusions presented on Page 3-5 are unsupported. For uranium, most of the individual results are less than 20 parts per million (ppm). Furthermore, all of the uranium data, except for a few points in certification units O-20, P19-23, and P18-40 of A1PI and in certification unit 3 of Area 8, Phase I (A8PI), were less than 40 ppm, which is less than half of the FRL. Only a single data point exceeded the FRL. Because essentially all results are far removed from the FRL, it is not surprising that decisions made based on different analytical methods are identical. As noted in original specific comment 3 above, most laboratory thorium 232 results are suspect. Therefore, few fully

viable data sets for that radionuclide are available for comparison. Finally, as noted in the review comments for the "Comparability of In Situ Gamma Spectrometry and Laboratory Data, Revision 1" report, the developed (radon-corrected) HPGe results for radium 226 are biased low relative to laboratory results. Such a bias, in conjunction with the tendency for HPGe results to have lower variability, would tend to result in the passing of certification units that would fail on the basis of laboratory results. The text should be revised to address these issues.

Response: DOE-FEMP disagrees with the comment. If remediation has been successful, all total uranium concentrations should be low and different analytical methods should all indicate low concentrations. The concentrations may or may not be comparable as defined in this report, but they should be low relative to FRLs and certification decisions can be made irrespective of the technique.

If the thorium data for A1PI are discarded, the comparability of HPGe and laboratory certification data will improve. A1PI thorium-232 data represent worse case scenarios from a comparability perspective.

Finally, radon corrected HPGe radium-226 results are not biased low relative to laboratory data. As shown in Table 8-5B of the January 1999 Comparability Study, the corrected HPGe radium-226 measurement is virtually identical with weighted laboratory data from the Field Quality Control Station. As shown in Table 3-5 of the January 1999 Certification Comparability Study, in four of the five data sets, the average corrected HPGe radium-226 data are higher than the average laboratory data.

Action: None.

RESPONSES TO U.S. EPA TECHNICAL REVIEW COMMENTS ON UPDATED SECTION 2.5 (REVISION B) FOR INCORPORATION INTO USER'S MANUAL, ENTITLED "CERTIFICATION" (JANUARY 1999)

FERNALD ENVIRONMENTAL MANAGEMENT PROJECT

GENERAL COMMENT

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 1
Comment: Because of the extent of comments and unresolved issues associated with HPGe comparability study documents, incorporation of updated Section 2.5 entitled "Certification" into the user's manual is not acceptable at this time. Issues related to HPGe comparability study documents should be discussed during Real-Time Workgroup meetings.

Response: Disagree with comment. DOE-FEMP believes that very few of the review comments are sufficiently substantive to preclude use of the HPGe for certification.

Action: None.

SPECIFIC COMMENTS

Commenting Organization: U.S. EPA Commentor: Saric
Figure #: 2.5-1 Page #: NA Line #: NA
Original Specific Comment #: 1
Comment: The figure discusses surveying sampling locations in the field to locate and finalize the given areas. However, it is not exactly clear what the text means by "survey." The text should be revised to clarify what type of survey will be conducted.

Response: Survey in this context is the process of delineating precise locations from coordinates. It does not refer to any type of monitoring activity. The reader can refer to Figure 3.9 and associated text in the Sitewide Excavation Plan for more details.

Action: None.

Commenting Organization: U.S. EPA Commentor: Saric
Figure #: 2.5-2 Page #: NA Line #: NA
Original Specific Comment #: 2
Comment: According to this figure, if a certification unit fails because of a large variation in the data, additional physical samples or HPGe measurements will be performed. However, if the unit still fails due to variability during further physical sampling or measurements, it is not clear what action would be taken. If the unit fails certification based on additional sampling data, either the certification unit should be further remediated or other DOE management decisions should be explored. The text should be revised accordingly.

Response: Figure 2.5-2 in proposed Section 2.5 of the User's Manual is a reproduction of Figure 3-10 of the Sitewide Excavation Plan. Issues dealing with physical samples should be addressed to the Sitewide Excavation Plan. With regard to HPGe, however, as noted in Section 3.7.3, 16 HPGe measurements would be taken. There are no plans to add additional HPGe measurements because of high variability.

Action: Clarify text on Lines 12-15 on Page 2.5-1 to note that no additional HPGe measurements will be added due to high variability.

RESPONSES TO U.S. EPA TECHNICAL REVIEW COMMENTS ON UPDATED SECTION 3.7 (REVISION B) FOR INCORPORATION INTO USER'S MANUAL, ENTITLED "CERTIFICATION MEASUREMENTS" (JANUARY 1999)

FERNALD ENVIRONMENTAL MANAGEMENT PROJECT

GENERAL COMMENT

Commenting Organization: U.S. EPA — Commentor: Saric
Section #: NA — Page #: NA — Line #: NA
Original General Comment #: 1

Comment: Because of the extent of comments and unresolved issues associated with HPGe comparability study documents, incorporation of updated Section 2.5 entitled "Certification" into the user's manual is not acceptable at this time. Issues related to HPGe comparability study documents should be discussed during Real-Time Workgroup meetings.

Response: See response to Original General Comment No. 1 on proposed Section 2.5 to User's Manual.

Action: None.

SPECIFIC COMMENTS

Commenting Organization: U.S. EPA — Commentor: Saric
Figure #: 2.5-2 — Page #: NA — Line #: NA
Original Specific Comment #: 1

Comment: According the this figure, if a certification unit fails because of a large variation in the data, additional samples or HPGe measurements will be taken. However, if the unit still fails due to variability upon further sampling or measurements, it is not clear what action would be taken. If the unit fails certification based on additional sampling data, either the unit should be further remediated or other DOE management decisions should be explored. The text should be revised accordingly.

Response: This comment is a repeat of Original Specific Comment No. 1 on proposed Section 2.5 of the User's Manual. Refer to the response to that comment.

Action: None.

Commenting Organization: U.S. EPA — Commentor: Saric
Section #: 3.7.5 — Page #: 3.7-3 — Line #: 23
Original Specific Comment #: 2

Comment: The guidance bullet states that if any HPGe detector result is a possible data outlier, it will be remeasured and an average of the two measurements (two out of two or the closer of two out of three) will be used for certification. No comparable remeasurement procedure exists for laboratory measurements; laboratory measurements are either rejected or considered valid and used as generated. The use of selected composite in situ measurements will tend to bias the overall statistics toward the mean. This will, in turn, decrease the probability that a unit will be rejected for certification. The use of such data

manipulation must be thoroughly justified. One alternative procedure would be to remeasure a certification unit using a new, randomly generated set of 16 HPGe measurement locations, when an apparent data outlier is detected in the original measurement set. The text should be revised accordingly.

Response: This bullet item is open to discussion and resolution in a future Technical Workgroup meeting. Its intent is to address possible environmental effects for which there are no analogs in laboratory measurements. In particular this bullet item is intended to address possible radon effects on radium-226 measurements.

Action: Discuss and resolve issue in a future Technical Workgroup meeting.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 3.7.5 Page #: 3.7-3 Line #: 24
Original Specific Comment #: 3

Comment: The intent of the second guidance bullet is unclear. It appears as though the specific criteria for HPGe measurements may lead to unnecessary field time for areas that would be considered clean. For example, applying the criteria established in the guidance bullet for the data associated with certification results of A8PI in the January 1999 comparability study would result in at least 14 additional measurements. Both HPGe measurements and laboratory analysis indicate that this unit meets FRLs. Furthermore, if these criteria are to be instituted, an RPD higher than 20 percent should be allowed for additional measurements. The text in this bullet should be re-evaluated and revised as necessary.

Response: Experience may show that this bullet item is unnecessary. Its intent was to account for possible environmental effects on HPGe data, particularly on radium-226 data. Applying the criteria to A8PI certification data would result in two additional measurements: One for total uranium in CU1 at Location 1-12-G and one for total uranium in CU3 at Location 3-14-G.

The review comment is correct in that if these criteria are instituted, a RPD greater than 20 percent for the second measurement would be required. In this regard, a RPD of 35 percent for the second measurement would be consistent with Section 5.4 of the ASL D Data Validation Checklist. Section 5.4 covers validation criteria for non-consecutive duplicates.

Action: Discuss and resolve bulleted guidance item in a future Technical Workgroup meeting.

RESPONSES TO U.S. EPA TECHNICAL REVIEW COMMENTS ON DRAFT DATA VALIDATION CHECKLIST FOR VALIDATING HIGH-PURITY GERMANIUM DETECTOR MEASUREMENTS TO ANALYTICAL SUPPORT LEVEL D (JANUARY 1999)

FERNALD ENVIRONMENTAL MANAGEMENT PROJECT

GENERAL COMMENTS

Commenting Organization: U.S. EPA — Commentor: Saric
Section #: NA — Page #: NA — Line #: NA
Original General Comment #: 1

Comment: As a result of Real-Time Workgroup meetings, it was generally understood that the ratio of low- to high-energy photon fluence would be evaluated during the data validation process. This step would provide a means for identifying a high-activity source located some distance underneath the soil surface. For this reason, the data validation checklist should be revised to include an element to account for data related to the ratio of low- to high-energy photon fluence.

Response: There is no need to include a separate element in the data validation checklist pertaining to the ratio of low-to-high energy photos. Item 1.3 on the data validation checklist refers to form FS-F-5509, which is the revised data review checklist that appears in Table 5.4.1 of the User's Manual. If unsatisfactory responses and explanations on that form are received for the low-to-high energy ratio items, data validation will stop per Item 1.3 until the deficiencies on the data review checklist are addressed.

Action: None.

Commenting Organization: U.S. EPA — Commentor: Saric
Section #: NA — Page #: NA — Line #: NA
Original General Comment #: 2

Comment: The attached "In-Situ HPGe Gamma Spectrometry Detector Annual Calibration Checklist" is generally deficient for supporting production of definitive data because it relies on multiple readings of point sources, followed by mathematical manipulation to generate a result. This sort of calibration is quite different from the use of the instrument to measure activity in a defined volume of space. This issue should be discussed during Real-Time Workgroup meetings.

Response: DOE-FEMP disagrees with the comment. The calibration methodology for in-situ gamma spectrometry detectors is approximately 30 years old and has been extensively peer reviewed via numerous articles appearing in scientific journals. See also response to Original General Comment No. 2 to the January 1999 HPGe Comparability Study. However, DOE-FEMP is willing to discuss the calibration issue at a future technical workgroup meeting.

Action: Discuss calibration method for in-situ gamma spectrometry at a future Technical Workgroup meeting.

SPECIFIC COMMENTS

Commenting Organization: U.S. EPA — Commentor: Saric
Section #: 1.1 — Page #: 1 — Line #: NA
Original Specific Comment #: 1
Comment: The data validator is instructed to determine the completeness of the data package. The checklist should either be revised to cite completeness criteria or provide a reference for these required elements.

Response: The data validation checklist is used by data validators. They do not specify completeness criteria, but rather validate to criteria spelled out elsewhere. The Real-Time Group is developing a procedure (34-00-005 entitled "In-Situ Gamma Spectrometry Data Validation") identifying the content of a data package. The completeness requirements will be spelled out in this procedure.

Action: Ensure completeness requirements are detailed in the Real-Time Group Data Validation Procedure 34-00-005 currently in preparation.

Commenting Organization: U.S. EPA — Commentor: Saric
Section #: 2.9 — Page #: 4 — Line #: NA
Original Specific Comment #: 2
Comment: The first phrase of this sentence is unclear. The phrase should be revised to read, "Were the daily checks (efficiency and resolution) performed with...." or whatever is meant.

Response: Agreed, the sentence is awkward. The intent of Section 2.9 is to ensure that the daily performance checks use the same standards (sources) as were used to generate the data to construct the control charts.

Action: Item 2.9 will be rewritten for clarity.

Commenting Organization: U.S. EPA — Commentor: Saric
Section #: 5.2 — Page #: 6 — Line #: NA
Original Specific Comment #: 3
Comment: The action items referred to here should be "...according to 5.3 or 5.5" and "...according to 5.4 and 5.6," respectively. This modification would provide explicit references to the low-activity procedures for evaluating duplicate field results.

Response: Agree with comment.

Action: The action items will be rewritten per the above comment.

Commenting Organization: U.S. EPA — Commentor: Saric
Section #: 5.4 — Page #: 7 — Line #: NA
Original Specific Comment #: 4
Comment: It is not clear why an allowance of 35 percent RPD is instituted for nonconsecutive duplicate results. Because the HPGe system averages radionuclide concentrations over relatively large areas, minor misalignments in repositioning the detector should not result in major disagreements in duplicate results. A major misalignment would be prevented by proper use of the GPS. Furthermore, certification measurements are intended to be

taken in areas considered to demonstrate homogeneous contamination. For [illegible] precertification HPGe measurements, a 35 percent RPD may be feasible; however, for certification, a 20 to 25 percent RPD would be more appropriate. This issue should be discussed during Real-Time Workgroup meetings and the text revised accordingly.

Response: Refer to response to Original General Comment No. 3 on the January 1997 Certification Comparability Report.

Action: Discuss issue of duplicates at a future Technical Workgroup meeting.

RESPONSES TO OHIO EPA COMMENTS ON THE "REVISED REAL TIME IN SITU GAMMA SPECTROMETRY REPORTS AND DOCUMENTATION TO SUPPORT THE USE OF HIGH PURITY GERMANIUM DETECTORS TO PERFORM FINAL SOIL CERTIFICATION FOR PRIMARY RADIONUCLIDES" PACKAGE

FERNALD ENVIRONMENTAL MANAGEMENT PROJECT

GENERAL COMMENT

Commenting Organization: Ohio EPA — Commentor: ODH
Section #: — Page #: — Line #: — Code: general
Comment #: 1

Comment: A support document, the In-Situ Gamma Addendum to the Sitewide CERCLA QAPP asserts under criterion 9 that management assessment of the real-time program shall occur annually in accordance with a plan. There is no mention of what this plan is. Is the addendum itself the plan?

Response: The comment refers to Criterion 9 in the document entitled "Real Time Instrumentation Measurement Program Quality Assurance Plan." Criterion 9 (Management Assessment) was placed in the QA Plan in order to be consistent with Document RM-0012 (FEMP Quality Assurance Program). Both Documents state that management assessments must be performed in accordance with a plan. These plans are not included in QA documents such as the two listed above, because they may vary from year to year, from manager to manager, or from program to program depending upon what management believes it is necessary to emphasize in a given assessment. For the Real Time Instrumentation Measurement Program (RTIMP), the manager of the RTIMP Group is responsible for developing the management assessment plan.

Action: None.

COMPARABILITY OF IN-SITU GAMMA SPECTROMETRY AND LABORATORY DATA, REVISION 1

Commenting Organization: Ohio EPA — Commentor: ODH
Section #: — Page #: — Line #: — Code: general
Comment #: 2

Comment: The revised In-Situ Spectroscopy Reports support the conclusion that HPGe detectors provide overall comparable decisions as discrete samples relative to soil certification decisions for primary radionuclides in certain areas at the FEMP. Thus far this has been demonstrated for "routine" geometries and locations from which the data was generated. Additional data should also be collected from areas characteristic of the entire suite of extreme topographic conditions likely to be encountered once excavation commences in the production area.

Response: Agree with comment. Toward this end DOE's Environmental Measurements Laboratory currently has recently completed a report on the use of HPGe in trenches. Section 4.9 (Topographic Effects) of the User's Manual is also helpful in assessing the effects of extreme geometries on HPGe data.

Action: Studies will be conducted on extreme geometries in the future as appropriate to demonstrate comparability between laboratory and HPGe measurements.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: General Comment Page #: NA Line #: Code: C
Comment #: 3
Comment: The use of HPGe for certification appears to require numerous input parameters such as soil moisture, radon concentration, and knowledge of the contaminant mix (heterogeneous vs. homogenous, etc), FCS measurements before and after sampling, etc. How many samples can be performed per day, and at what total cost as compared to taking physical samples? What is the cumulative effect of all the additional parameters on the measurement result?

Response: One way to estimate the total influence of all parameters on HPGe measurements is to look at the long term standard deviation. As discussed on Pages 2-8 and 2-10, the long term standard deviation is approximately 6 percent of the measurement value for total uranium and thorium-232 and about 10 percent of the measurement value for radium-226. These values are small relative to the intended use of the data and indicate that the contributions of mutiple parameters to the measurement uncertainty is within acceptable limits. Another way to assess the cumulative effect of all additional parameters is by examining the data in Table 3-6 of the January 1999 Certification Comparability Study. As noted in Section 3.4 of the same report, HPGe data generally have smaller percent relative standard deviations for total uranium and thorium-232 than do laboratory data, and hence are more representative for the purposes of certification. HPGe data generally have similar percent relative standard deviations as laboratory data for radium-226.

With regard to costs, personnel at Argonne National Laboratory developed detailed estimates of the cost effectiveness of using HPGe vs. laboratory analysis of physical samples. This study was given to OEPA. The Study showed that approximately $550,000 would be saved over the life of the project, if HPGe were used for certification rather than physical samples. Additionally, a detailed cost comparison of HPGe and physical samples is being made for certification measurements performed in A8PII. These will be made available to OEPA and U.S. EPA.

The number of HPGe measurements per day per detector is discussed in Section 4.13 of the User's Manual.

Action: Make HPGe and physical sample cost comparison data for A8PII available to OEPA and U.S. EPA.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: General Comment Page #: Line #: Code: C
Comment #: 4
Comment: The entire comparison has been performed using wet weight rather than dry weight, as is required for certification. The comparison between field moisture content and laboratory soil moisture content appear as if moisture content will add as much uncertainty to the measurement as any other factor.

Response: Depth variations in moisture content do add to the uncertainty in HPGe measurements. However, all of the data in the January 1999 Certification Comparability Study are presented on a dry weight basis. As shown in Tables 3-1 through 3-5 of that report and as concluded in Section 3.4 of that report, both HPGe and laboratory data result in overall comparability with respect to the 95 percent UCL of the mean being below the FRL for a given analyte. In turn, this means that the same certification decisions are made for total uranium, thorium-232, and radium-226 (with one exception; HPGe measured high radium-226 concentrations in one CU in A1PII due to high radon levels) regardless of whether HPGe or laboratory data are used. Further, taking into account the closeness of data to their MDCs, average HPGe certification data are very comparable to average laboratory certification data. In summary, the data in the January 1999 Certification Comparability Study indicate that the moisture content of soils is not a major factor with respect to the representativeness or comparability of HPGe data.

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 2.1.1 Page #: 2-1 Line #: 25-26 Code: C
Comment #: 5
Comment: The text states that measurements were taken twice daily..." unless ..other work priorities interfered." The QA/QC should require that an a.m. and p.m. measurement be conducted whenever the HPGe is used.

Response: There is no valid reason for taking a.m. and p.m measurements at the Field Quality Control Station for routine HPGe measurements. True, such measurements could provide information on radon accumulation near the soil surface, but adoption of the radon monitor technique (Section 5.3.2 of the User's Manual) provides better radon data. Further, the QC procedure (ADM-16) already specifies HPGe pre-operational and post-operation efficiency (accuracy) checks against certified standards. Thus, pre-operational and post-operational accuracy checks are already specified and additional measurements at the FCS would not contribute additional meaningful information.

Action: None.

Commenting Organization: Ohio EPA Commentor: ODH
Section #: 2.1.1 Page #: 2-2 Line #: 1-2 Code: C
Comment #: 6
Comment: The 20% downtime for any instrument appears high. How will this availability affect the implementation of HPGe in the field and what are the causes for the downtime. Reviewing the report, one gets the impression that the instruments worked without any problems.

Response: The text in 2.1.1 stated that of the five detectors used for daily HPGe FCS measurements, one detector is usually out of service. This statement was interpreted by the reviewer to mean that the HPGe detectors are "broken" 20 percent of the time. This is not the case.

Each HPGe analyzer unit is outfitted with a detector, computer, and associated electronics. If, for any reason, one of the components of a "unit" is not operating or not performing within specifications, the entire unit is taken out of service, tagged to be checked out, and

another unit taken to the field. Approximately 50 percent of the total 20 percent downtime results from system or component malfunction.

In addition, some of the being out of service time is due to a perfectly useable unit being routinely taken out of service for routine maintenance. This routine or preventative maintenance is an ongoing part of our maintenance program to prevent detector failures. Elements of the preventative maintenance are performed on a daily, weekly, quarterly, biannual, and annual basis depending upon the maintenance to be performed. The schedule for preventative maintenance is given in the *In-Situ* Gamma Spectrometry Maintenance/Preventative Maintenance procedure, EQT-37. About 50 percent of the total 20 percent downtime results from maintenance/calibration activities. In summary, the HPGe units are reliable as the above comment indicates, and experience to date indicates that the availability of HPGe units in the field has not been adversely impacted by downtime.

Action: None.

Commenting Organization: Ohio EPA Commentor: ODH
Section #: Page #: Line #: Code: general
Comment #: 7
Comment: What is the effect of soil moisture on measurement results? Please provide equation. The differences in measured moisture content are disturbing.

Response: Sections 3.8 and 5.2 of the User's Manual discuss moisture measurements. Equation 5 in Section 5.2 shows how soil moisture affects data calculations.

The FEMP intends to phase out the use of the Troxler soil moisture meter in favor of a soil moisture detector based upon reflection of near infra-red radiation. This instrument will hopefully eliminate the anomalous moisture readings noted in the review comment.

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 2.4 Page #: 2-4 Line #: 15 Code: C
Comment #: 8
Comment: During control measurement activity, is there any data showing as to whether that particular instrument was subsequently used in the field all day and then checked again? Also, is there a record of weather conditions in the field during instrument usage, including humidity?

Response: A field activity log shows the time and place each detector is used on a given day. This log also records any unusual (from the analyst perspective) weather, environmental, or field conditions. Humidity, the temperature, wind, etc can be obtained from the site meteorology station. As indicated in the ASL D data validation checklist and in HPGe operating procedures, each detector used on a given day must have both a pre- and post-operational efficiency check and energy calibration. Thus, each instrument is checked before and after use.

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 2-2C Page #: NA Line #: NA Code: C
Comment #: 9

Comment: Temperature appears to have a larger affect than reported. An extended analysis should be performed comparing measurements taken at less than 60 degrees F, and greater than 60 degrees F. Also, what is the manufacturer's recommended operating temperatures.

Response: The discussion in Section 2.4 and Figures 2-2C and 2-2E clearly show that the temperature effect is not large. Nonetheless, field crews must be sensitive to heat effects on HPGe system electronics. In this regard, the discussion below on temperature issues is taken verbatim from Section E.1.5 in the July 1997 Comparability Study.

The total temperature instability for the field-deployed HPGe system is typically <50 ppm per channel per degree C over an operating range of 0 degrees C to 50 degrees C, according to manufacturer's specifications. It should be noted that 50 ppm does not refer to concentration of analytes; it refers to drift in the assignment of channels in a multichannel analyzer to a gamma photon of a given energy. This means that a photopeak at the upper end of 8192 channels (HPGe) could drift as much as 20 channels ($50 \text{ ppm} \times 10^{-6} \times 8192 \text{ channels} \times 50$ degrees C) if the temperature went from 50 degrees C (122 degrees F) to 0 degrees C (32 degrees F). Table 5-6 indicates that the channel acceptance position criteria in the daily performance check is $\pm$ 3 channels at 1332.5keV. The 1332.5 gamma photon is assigned to channel 3553 in a multichannel analyzer. The temperature instability at this channel per degrees C is (50×10^{-6}) (3553) = 0.178 channels. Thus, to exceed the $\pm$ 3 channel QC acceptance criterion, a temperature shift of 3/0.178 = 16.8 degrees C (30.4 degrees F) is required. On days when a temperature change in excess of 30 degrees F occurs, field crews need to be especially aware of potential drift in the HPGe energy calibration. However, the normal operation of the field system employs a digital spectrum stabilizer to reduce spectrum drift. In addition, personnel have been trained to maintain spectrum energy calibration. Finally, the analysis software monitors the system response with respect to peak shift and photopeak resolution and performs analysis based on the actual spectrum. Thus, temperature instability in the multichannel analyzer does not appear to present operational limitations.

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 2-3C Page #: NA Line #: NA Code: C
Comment #: 10

Comment: There is a slight decreasing trend of total U vs. humidity, but this trend is not evident in Th-232 vs. humidity, why?

Response: The reason for the slight decreasing trend of total uranium vs humidity is not known. However, the effect is only on the order of about 3 ppm spanning a humidity range of nearly 75 percent. Thus, the effect is very small from a practical perspective.

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 2-3C Page #: NA Line #: NA Code: C
Comment #: 11
Comment: During this test, were the instruments used all day in the field between a.m. and p.m. measurements, or were the instruments used just for two readings. The instrument should have been used all day to identify any problems with electronics etc., from exposure to elevated humidity.

Response: Five different detectors were used over the course of a year to make the measurements shown in Figure 2-3C. In some cases a given detector was only used to make the morning and afternoon measurements (for example, on days when no other measurements were scheduled). In other cases the detector was used all day.

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: Figure 2-4C Page #: NA Line #: NA Code: C
Comment #: 12
Comment: The figure appears to indicate that measurements taken during colder weather biases the results low.

Response: Figures 2-4B and 2-4D, which are on a dry weight basis, indicate that measurements taken in colder months tend to be lower than those taken in warmer months. Or conversely, measurements taken in warmer months tend to be higher than those taken in colder months. As discussed in Sections 2.3 and 2.6, this effect is not believed to result from temperature differences, but rather from soil moisture gradients.

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: Figure 2-5C Page #: NA Line #: NA Code: C
Comment #: 13
Comment: A map indicating the location of the FCS relative to any onsite radiation sources should be included. Plotting radon concentrations with the Radium results would be beneficial.

Response: A map of the FCS location relative to potential on-site radiation sources is not necessary. Ratios of low to high energy gamma photons did not show evidence of shine.

Radon is in secular equilibrium with radium-226 in soil. Therefore, their soil concentrations are equal.

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 3.2.1 Page #: 3.1 Line #: 27-28 Code: C
Comment #: 14
Comment: The locations are listed as PBC-1 through PBC-10 and PBC-12 through PBC-19. Is there no location PBC-11?

Response: Location PBC-11 is the Field Quality Control Station, which is called RBS-11. 2300

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: Table 3-5 Page #: NA Line #: NA Code: C
Comment #: 15
Comment: The difference between the Troxler measurements and laboratory measurement for soil moisture content appear to vary substantially. During implementation in the field how will these differences be addressed?

Response: Although there is some variation between pairs of measurements, on average the difference is small (27.8 vs. 24.6). Table 3-5 indicates that Troxler measurements consistently yield moisture values higher than laboratory data. Part of this may result from loss of moisture during sample collection, transport to the laboratory, and preparation for laboratory moisture analysis. However, if Troxler measurements are truly biased high relative to laboratory soil moisture measurements, the net result would be higher dry weight concentrations. Thus, a degree of conservatism is built into HPGe measurements.

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 4.4 Page #: 4-5 Line #: 27-32 Code: C
Comment #: 16
Comment: A comparison between dry weights should also be done. The reporting concentration for certification is in dry weight.

Response: As explained in Section 4.4, the rationale in demonstrating comparability was to avoid possible heterogeneity effects in soil moisture within the field of view. In this regard, some areas such as PBC-01 and PBC-07 appear fairly homogeneous with regard to moisture. Other areas, such as PBC-18 and PBC-19, appear heterogeneous with respect to moisture. Correcting wet weight HPGe data to dry weight data involves the utilization of a single moisture value to adequately represent soil moisture within the field of view of the detector. Rather than introducing a possible additional source of error into the comparability assessment, the HPGe data were left on a wet weight basis.

The comment is correct in that reporting concentrations for certification are on a dry weight basis. That is why all data in the January 1999 Certification Comparability Study are reported on a dry weight basis. Note that in that report, the HPGe and laboratory data are comparable also (Section 3.2).

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 5.1 Page #: 5-1 Line #: 19-21 Code: C
Comment #: 17
Comment: Why was weighting factor for laboratory gamma spectrometry changed?

Response: The weighting factor was changed from one over the counting error to one over the counting error squared in order to be consistent with the methodology agreed to between OEPA and DOE-FEMP on how to calculate radionuclide concentrations utilizing multiple energy lines. Please refer to Letter DOE-0141-98, dated November 20, 1997, from Johnny Reising to Tom Schneider entitled, "Measurement and Calculation of Thorium-232."

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 7.6.2 Page #: 7-10,11 Line #: NA Code: C
Comment #: 18
Comment: These correction algorithms have been developed by specifically fitting the data. Are there any supporting theoretical equations that support/justify this algorithm?

Response: DOE-FEMP is not aware of theoretical equations or mathematical models for radon-radium equilibrium in soil. However, at a minimum, theoretical equations would be comprised of two components: 1) a kinetic component describing the kinetics of radon-222 generation from radium-226 and 2) a component describing the diffusion of radon-222 from the soil to the atmosphere. Additionally, the atmospheric buildup of radon near the soil surface under certain conditions would also have to be taken into account. The solution to such equations might have to be numerical rather than analytical. In any event, the solution to the equations would not have the same form as the correction algorithm.

Lines 7 through 9 on Page 7-12 are reproduced below and summarize well the validity of the correction algorithm:

> "The radium correction factor that has been developed is empirical and hence is only valid for use at Fernald. It is a site-specific correction factor and should not be used universally. Other sites may use the methodology outlined here to develop their own site-specific correction factors."

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 7.6.2 Page #: 7-12 Line #: 7-9 Code: C
Comment #: 19
Comment: The limited application of the correction algorithm and presence of the K65 silos at the FEMP suggest that the correction for Ra-226 may not be applicable to the entire FEMP site. The document should state that where the correction factor has been applied, there is evidence that the correction factor may be applicable to the site.

Response: The correction algorithm is probably applicable to most, or all, of the site because the soil type is reasonably uniform over the site. The presence of the K65 silos is an entirely different issue. The concern about the silos is the generation of radon which could subsequently settle near the ground surface giving anomalously high "apparent" radium-226 corrections. This effect is compensated for by use of the radon monitor as described in Section 5.3.2 of the User's Manual.

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 8.0 Page #: General Comment Line #: NA Code: C
Comment #: 20

Comment: A comparison between FEMP HPGe and DOE EML HPGe is of limited value due to the extensive consulting provided by DOE EML to FEMP during the development of FEMP HPGe program.

Response: DOE-FEMP disagrees with the comment. True, DOE-EML has consulted extensively with the FEMP in the development of the FEMP HPGe program. However, the measurements were taken with different field crews, using different instruments, with different operating and calibration procedures, and utilizing different software packages. Further, EML has a quality control oversight role for radiochemistry measurements within the DOE-complex. As such, their measurements, whether field or laboratory, are the standard against which other measurements must be judged. For these reasons the comparison of DOE and FEMP data are valid.

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 8.2.2 Page #: 8-2 Line #: 20-24 Code: C
Comment #: 21

Comment: The text states that Wilcoxon Signed Ranks Test indicating bias between DOE-EML data and FEMP data, but offers no explanation for possible biases. Please provide explanation.

Response: DOE-FEMP does not have an explanation for the high biases of FEMP 31 cm data for total uranium and thorium-232 relative to EML data. However, as shown in Table 8-1 the biases are small, and as stated in Section 8.2.2 the 100 cm data are not biased. Further, despite the small biases noted above, the percent RPDs between EML and FEMP data meet the criterion for very good comparability (Section 8.2.2).

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 8.3 Page #: 8-3 Line #: 19-24 Code: C
Comment #: 22

Comment: This entire paragraph should be deleted. The conclusion that R. T. Reiman's results and EML's results are representative of what the FEMP would have measured is pure conjecture. FEMP should have participated.

Response: Agree with comment.

Action: The paragraph in question will be deleted as written.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 8.3 Page #: 8-3 Line #: 4-5 Code: C
Comment #: 23

Comment: Earlier in the text, it was stated that FEMP no longer uses EGAS software, but here it states that FEMP uses it. This is inconsistent.

Response: Note the parenthetical on Line 5 of Page 8-3 that states "GammaVision" has been customized to yield as similar results as possible to EGAS." Extensive testing of the two software packages was carried out in order to ensure continuity of data. The comment is technically correct, however, and Lines 4-5 will be reworded to resolve the inconsistency.

Action: Rewrite Lines 4-5 to clarify use of EGAS and GammaVision.

Commenting Organization: Ohio EPA Commentor: ODH
Section #: Page #: G-2 Line #: NA Code: C
Comment #: 24
Comment: The nuclide library list the 1001-0 keV gamma as originating from Th-234. This appears to be a typo as this gamma is emitted from Pa-234m.

Response: Agree with comment.

Action: The table will be corrected to indicate that the 1001.0 keV gamma is emitted from Pa-234.

COMMENTS ON COMPARABILITY OF IN SITU GAMMA SPECTROSCOPY AND LABORATORY DATA AND DECISIONS FOR CERTIFICATION UNITS, REVISION 0

The Ohio EPA has no comments.

COMMENTS ON RTRAK APPLICABILITY STUDY, REVISION 2

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 1.2 Page #: 1-3 Line #: 1-3 Code: C
Comment #: 25
Comment: The GPS may be able to display the speed at which the operator of the RSS is walking, aiding in the consistency of measurements.

Response: Agree with comment. The GPS is used to guide the operator to walk at 1 mph (1.47 feet per second).

Action: Change text in Section 1.2, Page 1-2, Lines 26 through Page 1-3, Line 3 to read:

> "The travel speed of the RTRAK can be continuously monitored by either the tractor's speedometer or the onboard GPS. Therefore, the speed of the RTRAK can be maintained reasonably well, and the field of view is consistent. The travel speed of the RSS is continuously monitored only by the onboard GPS. Any variations in the RSS field of view are attributable to the difficulty in attempting to maintain a constant walking speed over irregular terrain. The variations in the RSS field of view do not create serious difficulties in evaluating the data because the RSS is used to monitor small areas."

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 1.4 Page #: 1-4 Line #: 18 Code: C
Comment #: 26
Comment: In this section and throughout the document the total uranium measurements are based on "normally" enriched uranium. OEPA assumes this to mean "natural" abundance's of the uranium isotopes. Are there any checks or guidelines to ensure this assumption is true, especially when surveys are performed in the former production area.

Response: RI/FS Reports and the Record of Decision for OU5 conclude that the assumption of soil contamination by "normally" enriched uranium is valid, and any deviations from normality would be small and negligible from a risk perspective.

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 2.2 Page #: General Line #: General Code: C
Comment #: 27
Comment: The calibration of the RTRAK is based on static measurements from the HPGe and the RTRAK. These detectors have different fields of view. How is this accounted for in the calibration process.

Response: In areas for which radionuclides are uniformly distributed, matching the fields of view of the HPGe or NaI detectors for calibration is not an issue since the results should be the same independent of the field of view that is used. However, if the radionuclides are not uniformly distributed, then three geometric factors may result in inconsistencies between HPGe and RTRAK measurements: (1) any difference in the positions of the HPGe and NaI detectors when measurements are made, (2) differences in the sizes of the fields of view of the detectors, and (3) differences in the shapes of the fields of view of the detectors. The NaI detector does not have a circularly shaped field of view. Additionally, the orientation of the NaI detector (which is determined by the orientation of the NaI detector on the RTRAK) relative to any pattern in the distribution of the radionuclides may have some effect on RTRAK results. Areas with a significant amount of heterogeneity in the distribution of radionuclides were not used for calibration of the RTRAK. However, some small amount of heterogeneity will always be present that will affect the agreement between HPGe and RTRAK results due to the three factors listed above. Therefore, achieving the best fit between RTRAK and HPGe results was used as the basis for establishing the HPGe detector height to be used when calibrating the RTRAK. Regressions were carried out using data obtained for HPGe detector heights of 15 cm, 31 cm, and 1 m. The best results were obtained using 31-cm data. Statistically, 31-cm HPGe data provide the best match with the RTRAK results for the data set used for calibration. Therefore, calibration using a 31-cm HPGe detector height is preferred. Uncertainties associated with any less than perfect agreement between HPGe and RTRAK results, including any differences related to fields of view, are reflected in the uncertainties in the parameters in the calibration equations.

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 2.3 Page #: 2-6 Line #: 11-14 Code: C
Comment #: 28
Comment: This paragraph states that 1 mph/4 sec acquisition time might be the optimal operating conditions for the RTRAK. Besides the logistical considerations, how was this proven to be the best speed/acquisition time for the RTRAK?

Response: The 1.0 mph/4 sec acquisition time is a trade off between good precision and logistical considerations. Precision improves with longer data acquisition time and slower speeds while logistical considerations deteriorate with longer data acquisition times and slower speeds. This trade off is delineated in general conclusion 1 on Page 6-3 of Section 6.0 (Summary and Conclusions).

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 3.1 Page #: 3-1 Line #: 1-9 Code: C
Comment #: 29
Comment: This paragraph should explicitly define the ranges for the detections of contaminated soils. The depths and areal extent of contamination which can be detected utilizing the RTRAK.

Response: Sections 4.1, 4.2, 4.3 and Figure 3.3-2 of the User's Manual address the depths and areal extent of contamination which can be detected using RTRAK. These sections will be summarized and referenced in Section 3.0.

Action: Add a summary to Section 3.0 as described above.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 3.3 Page #: 3-2 Line #: Code: C
Comment #: 30
Comment: What are the accuracy limits of the GPS, and how would inexact locations affect the efficiency calibration of the RTRAK?

Response: GPS is thoroughly discussed in Section 5.8 of the User's Manual. As indicated in Section 5.8.1, horizontal RMS error as small as 15 cm can be achieved. The efficiency calibration of the RTRAK is independent of GPS errors. Both RTRAK and HPGe measurements were made at the exact locations as determined visually on the same day.

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 3.3 Page #: 3-3 Line #: 24 Code: C
Comment #: 31
Comment: This equation implies that the efficiency for measuring Th-232 is in the range of 6%. This efficiency is not clearly carried into MDC calculations. Provide information on how efficiency is used in the MDC calculations.

Response: Section 4.3.2 discusses location specific MDCs. The *a posteriori* MDCs can be calculated by applying standard propagation of uncertainty relationships to the RTRAK calibration equations and adding uncertainty contributions from non-counting sources. To calculate the MDC for a measurement, the assumption is made that the gross number of counts is equal to the number of background counts. The contributions of the non-counting uncertainty sources are set equal to the values used in calculating the MDCs in Section 4.3.1. The relationships on the bottom of Page 4-28 can be used to calculate the *a posteriori* MDC.

Action: None.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 4 Page #: Table 4-12 Line #: Code: C
Comment #: 32
Comment: The table should include the area over which the measurement is aggregated as well as the minimal detectable "hot spot" area and concentration.

Response: The first column in Table 4-12 gives the radius of the circle in feet over which the measurements are aggregated. The area can easily be calculated from these values. The "radius" heading in the first column was inadvertently deleted for the 2 sec/2 mph table but was placed in the 8 sec/0.5 mph table.

The minimal detectable hot spot area and associated concentration constitutes a table in and of itself, and cannot be included in the format of Table 4-12. This information, however, is contained in Figure 3.3-2 of the User's Manual.

Action: Insert the heading "Radius (ft)" in column one of Table 4-12 for 2 sec/2 mph operating conditions.

Commenting Organization: Ohio EPA Commentor: OFFO
Section #: 6.0 Page #: 6-4 Line #: 4-5 Code: E
Comment #: 33
Comment: This sentence does not make sense as written, believe it should read, "The current calibration equations provide good agreement with HPGe except at low uranium concentrations."

Response: Agree with comment.

Action: Rewrite sentence as suggested in the above comment.

COMMENTS ON DATA VALIDATION CHECK LIST FOR HPGe AT ASL LEVEL D, DRAFT

Commenting Organization: Ohio EPA Commentor: ODH
Section #: 5.7 Page #: Line #: Code: C
Comment #: 34
Comment: A field control station located north of the STP has been used as part of the QC program for both the RTRAK and HPGe. What provisions are in place to replace this locale and its function upon remediation of the former area?

Response: DOE-FEMP ultimately plans to use the calibration pad, which will be constructed in August 1999, to replace the field quality control station. We have also selected an alternative location for a second FCS near the North Access Road near RIMIA and are currently taking measurements there routinely.

Action: None.

COMMENTS ON UPDATED SECTION 2.5 (REVISION B) OF USER'S MANUAL, ENTITLED "CERTIFICATION"

The Ohio EPA has no comments on this.

COMMENTS ON UPDATED SECTION 3.7 (REVISION B) OF USER'S MANUAL, ENTITLED "CERTIFICATION MEASUREMENTS"

Commenting Organization: Ohio EPA Commentor: ODH
Section #: 3.7.5 Page #: Line #: Code: C
Comment #: 35

Comment: HPGe certification decisions for radium-226 appear equivalent to lab data provided radon-222 disequilibrium in soil and accumulation near the ground surface are compensated for as needed. The process for conducting these particular measurements and data review guidance from this section of the User's Manual make this process seem difficult to implement efficiently from a field operations viewpoint.

Response: The process for utilizing the radon monitor to compensate for radon accumulation near the ground surface is not difficult to implement inasmuch as it involves just setting up an HPGe detector in an area in close proximity to the area being measured so that it makes 15 minute counts continuously every 15 minutes. However, utilizing the radon monitor does result in some operations inefficiency. The data review process for radium-226 also requires some extra effort, however all calculations are performed electronically. Experience has shown that so far neither the utilization of the radon monitor nor the data review process has significantly affected daily operations.

Action: None.

CROSSWALK BETWEEN U.S. EPA COMMENTS ON COMPARABILITY STUDY DOCUMENTS AND CROSSWALK BETWEEN U.S. EPA COMMENTS ON RTRAK APPLICABILITY STUDY

The Ohio EPA has no comments on these two submittals.

RESPONSES TO U.S. EPA TECHNICAL REVIEW COMMENTS ON REVISED REAL-TIME IN SITU GAMMA SPECTROMETRY REPORTS AND DOCUMENTATION TO SUPPORT THE USE OF HIGH-PURITY GERMANIUM DETECTORS TO PERFORM FINAL SOIL CERTIFICATION FOR PRIMARY RADIONUCLIDES (ENCLOSURE 2)

FERNALD ENVIRONMENTAL MANAGEMENT PROJECT

GENERAL COMMENTS

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: NA (Not Applicable) (general comments on statistical methods)
Original General Comment #: 1
Comment: There are many methods available for comparing two sets of data. This report compares the results of *in-situ* gamma spectrometry measurements with laboratory data by:

1. using the average absolute percent difference between the certification unit (CU) means for each contaminant derived from each set of data, and

2. comparing the results of the compliance determination that would be made based on the upper confidence level (UCL) for the CU mean indicated by each set of measurements.

Results of the first comparison demonstrate that the two sets of measurements are almost always within 35% RPD. Exceptions are identified that indicate large values for the RPD occur only when the results are near the detection limit of the methods.

Results of the second comparison show that almost all CUs demonstrate compliance with the required remediation levels regardless of the data set used. Exceptions generally occur when unusually high laboratory measurements are observed without corroboration from the *in-situ* data. The *in-situ* method integrates over a much wider area than is covered by the collocated soil sample. The laboratory analysis of the soil samples is only expected to produce higher results than the *in-situ* method when small, isolated areas of elevated activity occur within the *in-situ* field of view and are included as part of the soil sample.

MARSSIM and EPA QA/G-9 (*Guidance for Data Quality Assessment*) provide guidance on evaluating data. Part of this evaluation is selecting the appropriate statistical tests. The basis for selecting these tests should be discussed, and the discussion should describe how these tests meet the objectives for the survey.

Both methods consider the mean values obtained in each CU. However, page 2-5 lists the mean, the 95 percent UCL, and the maximum concentration as considerations for compliance demonstration. There are additional comparisons that should be considered based on individual data points for the co-located data including correlation, regression, and graphical methods.

Correlation and regression analysis can identify the degree of association between co-located data. Outliers that affect the results may be clearly identified and subject to further investigation. Graphical analysis would better summarize the large volume of data presented in this report. Scatter plots comparing the results of the two analytical methods would demonstrate the degree of correlation, or lack thereof. A fundamental tool for comparison of alternative decision-making methods is the operating characteristic curve (OCC). An empirical estimate of the curve can be constructed by plotting the cumulative rejection percentage on the vertical axis and the contamination level on the horizontal axis. A good decision tool will produce an OCC that remains relatively flat at concentration far below the remediation level, rises sharply as the remediation level is approached, and approaches 100% at contamination levels slightly above the remediation level. OCCs may be constructed for data sets containing either co-located or non-co-located data points, and may include 1 or more CUs providing a broader range for the comparison. The empirical OCC permits a rapid visual determination of three important features of each measurement method:

1. ability to maintain a low false positive rate (Type II error) at contamination levels far below the remediation level,

2. ability to maintain a low false negative rate (Type I error) at contamination levels above the remediation level, and

3. a sharp rise in the rejection rate as the remediation level is approached

An important consideration is the application of the results of the comparison to other areas at FEMP. The current analysis is not constructed to facilitate this consideration. The current analysis reports only the average mis-classification rates averaged over all contamination levels, and the results are inconclusive. The mis-classification rates must be reported as a function of the contaminant concentration level to permit extrapolation to future surveys. If the areas and CUs where these methods might be applied in the future contain concentration higher than those encountered in this study, the current analysis is inadequate to predict the performance of the two measurement methods for making decisions regarding compliance in these areas. Constructing an OCC will help improve the comparison and permit extrapolation of the results to other areas.

Response: The comment is correct in that alternative statistical methodologies exist to compare data sets. However, for consistency with other site documents the data were evaluated in two ways: 1) as specified in the Sitewide Excavation Plan (SEP) for ascertaining that remediated soil meets certification criteria, and 2) as outlined in the January 1999 HPGe Comparability Study in which comparability criteria are delineated.

Action: None.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: NA
Original General Comment #: 2
Comment: The majority of the time the decision made based on *in situ* gamma spectrometry results and laboratory data on whether or not the release criterion has been exceeded is exactly the same. In the few instances where there is a disagreement, the average is below the

final remediation level and the difference in the uncertainty causes the disagreement. The comparison based on the 95th upper confidence level is consistent with Superfund guidance. However, this comparison only addresses the Type I decision error where a contaminated site may be incorrectly released. Using hypothesis testing (as recommended in MARSSIM, Superfund's Soil Screening Guidance, and EPA's QAD guidance documents) both Type I and Type II decision errors can be addressed. Evaluating the Type II decision error can help determine if the reason for these differences is based on an insufficient number of measurements (not enough power) or if the level of contamination is a problem. Also, the overall survey design can be evaluated to determine if it can be used effectively during subsequent decommissioning activities at the site.

Response: In the few instances where disagreement occurs in the certification decisions as to whether the release criteria has been exceeded or not, the basis of disagreement is clear. Situations where HPGe data for radium-226 exceed the 95 percent UCL result from high radon-222 concentrations. Situations where laboratory data for thorium-232 exceed the 95 percent UCL result from faulty laboratory methodology.

Thus, use of hypothesis testing to evaluate Type II errors does not appear to be necessary. With respect to overall survey design, the purpose of this report is not to evaluate the approach to certification, but rather to compare decisions. The approach to certification is outlined in the SEP.

Action: None.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: NA
Original General Comment #: 3

Comment: There is no discussion of the development of data quality objectives (DQOs) for this project and how they might affect the interpretation of results. The development of the final remediation levels is not discussed. Have the certification unit boundaries been compared to the assumptions used to develop the final remediation levels to ensure they are compatible? What other assumptions were used to develop the survey design? The level of detail provided with this report is insufficient to evaluate the overall demonstration of compliance. Only the statistical test results are provided, and only those results have been reviewed.

Response: It is beyond the scope of this report to present DQOs for certification and to address other issues raised in the comment. DQOs were developed for physical sample collection and analysis. DQOs were also developed for HPGe certification measurements. However, the intent of this report is to demonstrate comparability of measurements taken in a given area with one analytical technique to measurements made on physical samples in the same area with a second analytical technique. The intent of this report is not to demonstrate compliance with final remediation levels. That is the purpose of certification reports. Other issues raised in the comment are discussed in the SEP.

Action: None.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: Section 3
Original General Comment #: 4

Comment: Section 3 discusses the comparability of the results. The methods used to compare the data include the assumption the data are normally distributed. There is no discussion of whether or not this assumption has been verified. The use of nonparametric statistics is recommended in MARSSIM and Superfund's Soil Screening Guidance to avoid verifying assumptions of normality. Graphically representing the results of a preliminary data review (as recommended in MARSSIM and EPA QA/G-9) would improve the readability of this section. The use of posting plots, histograms, and power curves could be used to visually interpret the results, compare the two methods, and evaluate the effectiveness of the survey design.

Response: The statistical analysis of certification samples at the FEMP follows guidelines in the SEP (Appendix G, Table G-2 provides the selection procedure for the certification statistical analysis method). The methodology as outlined in the SEP for evaluation of the effectiveness of the survey design is consistent with MARSSIM and has been approved by the U.S. EPA. The methodology includes a test for normality, log-normality, and the alternative application of non-parametric tests. For actual certification documents (i.e., certification reports) this approach is followed.

The Area 1, Phase I (A1PI) laboratory samples-based certification decisions were determined following the guidelines outlined in the SEP. However, the comparability study-based HPGe measurement results performed in the A1PI CUs were not evaluated using the methodology outlined in the SEP.

Action: Perform the necessary data evaluations, as outlined in the SEP, on the A1PI Comparability Study Part A HPGe measurement results.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: NA
Original General Comment #: 5

Comment: The references (page R-1) are not all referred to in the text. For those that are, it is difficult to determine which reference is being cited. The text makes extensive references to other reports with the assumption that the reader has access to these reports. For example, page 2-1, 4th paragraph, last sentence; page 2-4, beginning sentences of Section 2.2.2; page 2-5, 1st sentence following the definition of relative percent deviation (RPD, also referred to as percent relative deviation).

Response: Agree with comment.

Action: The references on Page R-1 will be clarified to make it easier for the reader to ascertain which reference is being cited.

SPECIFIC COMMENTS

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 2.0
Original Specific Comment #: 1
Comment: The concept of the relative standard deviation (RSD, or coefficient of variation) should be defined in this section. The RSD is used as part of the data evaluation in Section 3.0.

Response: Agree with comment.

Action: The relative standard deviation will be defined in Section 2.2.2.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Figures #: 2.0 Page #: 2-5
Original Specific Comment #: 2
Comment: Page 2 through 5 notes that a greater number of measurements or CUs will increase the representativeness of each data set and reduce variability. However, it is not clear if the contaminants are assumed to be homogeneously distributed or if the presence of small areas of elevated activity is suspected. If small areas of elevated activity are present, increasing the number of measurements may result in an increase in variability.

Response: Agree with comment. If multiple small areas of high activity are present within a given CU, then increasing the number of measurements may result in an increase in variability. However, such areas are unlikely to be present because of efforts during precertification to locate any areas of elevated contamination.

Action: A sentence will be added at the end of Line 16 on Page 2-4 to note that variability could be increased by increasing the number of measurements within a given CU.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 2.1
Original Specific Comment #: 3
Comment: The selection of measurement methods and sample locations does not address whether the areas and soils reflect different types of materials (e.g., native soils and imported backfill). The text should describe surface conditions such as ruts, furrows, undulations, etc., which might impact the *in-situ* measurements by modifying the detector's field of view. There is no discussion of soil properties that might lead to elevated radon accumulations at or near the surface. The characterization of soil properties should discussed in more detail along with the potential impact on the analytical results and how this might affect the variability in the results.

Response: The issues noted in the comment above are generally addressed in the User's Manual. Section 4.9 discusses topographic effects. Section 5.3 discusses the use of the radon monitor for radon corrections. Section 5.4 discusses the review of data. Any anomalous situations encountered in measurement of the data would have been noted in data review and discussed in the text as appropriate.

Action: None.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 2.2.1
Original Specific Comment #: 4
Comment: The goal stated in the first sentence does not mention small areas of elevated activity as a potential concern. However, the third bullet includes the maximum concentration as a consideration for demonstrating compliance, which implies that small areas of elevated activity might be a concern. MARSSIM guidance describes an integrated survey design that combines concerns of homogeneous and heterogeneous contaminant concentrations into a single survey design. Also, the use of the term "hot spot" is not recommended because it implies there is automatically a health risk when this may or may not be the case.

Response: The basis of the survey design used at the FEMP has been detailed in the SEP and approved by the U.S. EPA. It is similar to MARSSIM but not identical. Importantly, though, the survey design includes a complete scan in precertification which addresses the issue of potential hot spots. The term "hot spot" is defined in the SEP and is used at the FEMP within the context of that definition.

Action: None.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 2.2.2
Original Specific Comment #: 5
Comment: A brief summary of the comparability discussion between HPGe and laboratory data would be helpful for readers who may not be familiar with the December 1998 Comparability Study.

Response: It was not anticipated that this document would be reviewed on a stand-alone basis by reviewers not familiar with the January 1999 (December 1998) Comparability Study. An appendix will be added to this report which contains the executive summary of the January 1999 (December 1998) Comparability Study.

Action: Add appendix which contains the executive summary of the January 1999 Comparability Study.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 2.3
Original Specific Comment #: 6
Comment: The role of "hot spot criteria" in comparability should be discussed. This is particularly important for the *in-situ* measurements since they represent the "average" activity over a large area. A CU that demonstrates compliance using the *in-situ* results may fail based on the "hot spot" criteria. MARSSIM provides guidance on evaluating the dose or risk from a small area of elevated activity based on the size of the area and the radionuclide-specific relationship between area and dose or risk.

Response: The comment is correct. However, "hot spots" are not an issue for these particular data sets, and thus are not addressed. The reviewer is directed to the SEP for a detailed discussion of "hot spots."

Action: None.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 3
Original Specific Comment #: 7
Comment: The data listed in Table 3-5 and Table 3-6 have not been corrected for the different numbers of measurements from *in-situ* and laboratory analyses. For example, there are 457 laboratory results and 399 *in-situ* results for the A1PI 32-cm and 100-cm data sets. The analysis based on the RPD does not take the different number of results into account. In assessing the variability of the data, the standard error (standard deviation divided by the square root of the number of measurements) might be considered which takes into account the different number of data points.

Response: The assumption is that because the number of samples in the data sets are large that the RPD would not be affected by sample size. The use of RPD does not constitute a vigorous statistical test, but rather a practical way of comparing this data in the same terms defined for comparability of two measurements in the January 1999 Comparability Study.

Action: None.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 3.1
Original Specific Comment #: 8
Comment: The discussion addressing comparability of total uranium should clearly state the assumption about the isotopic distribution. This is especially important for a site like Fernald where different mixtures of uranium were used.

The UCL for nine CUs were artifacts associated with the laboratory procedure, which is described in a separate document. The discussion should include a brief explanation of the important details and assess their significance to the analysis.

The discussion of anomalous conditions leading to high near-surface radon accumulations should include the rationale for identifying these results as anomalous instead of accurate representations of the ^{226}Ra concentrations. Given the significance of this finding and the implications in correcting for radon equilibrium, there is a need to present more information on the nature or condition of the soil that could have produced these results. There should be an assessment of how prevalent this condition might be at Fernald (e.g., is this an isolated incident) and how this information can be used to identify similar conditions at other locations.

Response: The comment is technically correct about discussing assumptions and giving brief explanations. However, this document was not intended to be a stand-alone report to be reviewed by persons unfamiliar with the Fernald site. The assumption was that the document would be reviewed by members of the Technical Workgroup who are familiar with much/most of the supporting information and documents. Consequently, all of the background, assumptions, and explanations that might normally accompany a stand-alone report were omitted in the interest of brevity.

Action: None.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 3.2
Original Specific Comment #: 9
Comment: The basis for selecting the 20% RPD and 35% RPD action levels for comparability should be identified and explained. EPA QA/G-5 (*EPA Guidance for Quality Assurance Project Plans*) includes a discussion of evaluating comparability as a survey objective. For this analysis selecting a single value rather than a range of values may provide a criterion that is more technically defensible, or at least easier to defend. Also, a discussion concerning changes made in the way laboratory results were calculated and why this resulted in lower ^{232}Th concentrations should be included.

Response: The basis for selecting 20 percent RPD and 35 percent RPD action levels for comparability is presented in Section 4.0 of the January 1999 Comparability Study. The discussion of the way in which laboratory results were calculated to result in low Thorium-232 concentrations is outlined in a series of memoranda to the U.S. EPA and is also explained in the A1PI Certification Report. It is beyond the scope of this report to reproduce this material.

Action: Reference the A1PI Certification Report as a document in which an interested reader could obtain additional information relative to the laboratory analysis of Thorium-232.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 3.3
Original Specific Comment #: 10
Comment: The percent relative standard deviation (RSD) reported in Table 3-6 do not use the same evaluation criteria used for RPD. The discussion of representativeness should identify a criterion for deciding whether or not the results are acceptable. EPA QA/G-5 (*EPA Guidance for Quality Assurance Project Plans*) includes a discussion of evaluating representativeness as a survey objective.

This section should include a discussion about the interpretation of the results and their applicability to other areas of the FEMP. See the general statistics comments for additional comments.

Response: Evaluation criteria have not been specified for the percent relative standard deviation. It is not used as a comparability parameter. Instead, it is used in a qualitative sense only to show that generally the percent RSDs are smaller, as expected, for HPGe data than for laboratory data.

A discussion of the general applicability of these data to other areas of the FEMP was deemed to be premature given the sensitive nature of use of the HPGe for certification. Such discussions were thought to be better addressed in Technical Workgroup meetings.

Action: None.